logged 4-5-72

W R Hudson

HIGHWAY RESEARCH RECORD

6

Number 369

Choice of Travel Mode and Considerations in Travel Forecasting

19 Reports

HIGHWAY RESEARCH BOARD

NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING

The opinions and conclusions expressed in this publication are those of the authors and not necessarily those of the Highway Research Board.

HIGHWAY RESEARCH RECORD

Number 369

Choice of Travel Mode and Considerations in Travel Forecasting

19 Reports

Subject Areas

55 Traffic Measurements
84 Urban Transportation Systems

HIGHWAY RESEARCH BOARD

DIVISION OF ENGINEERING NATIONAL RESEARCH COUNCIL
NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING

WASHINGTON, D.C. 1971

ISBN 0-309-01986-9

Price: $6.00

Available from

Highway Research Board
National Academy of Sciences
2101 Constitution Avenue
Washington, D.C. 20418

CONTENTS

FOREWORD

Inasmuch as state departments of transportation have now become a reality and new transportation modes and services are being actively considered in urban areas, the need for accurate travel demand forecasting is greater than ever before. Transportation services of all kinds will need to be improved to meet ever-increasing travel demands. Forecasts of these demands, by mode of travel, are essential to good decision-making.

Papers in this RECORD give examples of methods for forecasting high-speed rail use, air travel demands, and the potential use of interurban and intraurban transit.

How much service should be provided? Where? By what mode? What will be the effect on other modes? These and other questions must be answered before projects can move to the implementation stage.

Although existing techniques and models (used primarily for highway travel forecasting) can be adapted to these expanded requirements, many qualified observers feel that new methods need to be developed to respond to these issues. New models based on observations of traveler behavior can help provide the information needed to estimate the potential of new modes of travel. Criticism of the existing (and sometimes cumbersome) transportation forecasting models developed in the last decade is mounting. Cannot direct estimates of travel demands for a particular mode or facility be made? Must all forecasts of future use consider all trip and all modes, even those that potentially will not be served by the facility? Then, too, present techniques rely heavily on observed behavior conditioned by existing, not future, transportation systems. Can we safely extrapolate such behavior by using models that are calibrated to the available data base? Several papers in this RECORD deal with this subject and present new approaches to travel forecasting. Others concentrate on improving techniques already in use.

We stand at a crossroads in the development of travel forecasting methodology. Shall we discard what has been done and opt for new, perhaps more promising, methods and techniques? Or should we continue to use existing methods because we are familiar and comfortable with them? The answer is, of course, that new methods that prove their worth through use are always replacing older ones. Let us give these new methods that opportunity, even as we constantly improve present techniques.

—George V. Wickstrom

SPONSORSHIP OF THIS RECORD

GROUP 1—TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
J. Douglas Carroll, Jr., Tri-State Transportation Commission, chairman

<u>Committee on Transportation Forecasting</u>
George V. Wickstrom, Metropolitan Washington Council of Governments, chairman
Mark M. Akins, Daniel Brand, Austin E. Brant, Jr., Henry W. Bruck, Francis E. Coleman, Thomas B. Deen, William L. Grecco, Charles H. Groves, John R. Hamburg, Lawrence V. Hammel, Donald M. Hill, Wolfgang S. Homberger, Frank E. Horton, G. H. Johnston, Louis E. Keefer, Robert Kochanowski, Herbert S. Levenson, Brian V. Martin, James J. McDonnell, Eugene R. Mills, Kenneth W. Shiatte, Paul W. Shuldiner, Vergil G. Stover, Anthony R. Tomazinis, Robert T. Wood, Richard D. Worrall

<u>Task Force on Value of Travel Time</u>
Thomas E. Lisco, Chicago Area Transportation Study, chairman
Neil W. Mansfield, David Quarmby, Shalom Reichman, Peter R. Stopher, Thomas C. Thomas

James A. Scott and Kenneth E. Cook, Highway Research Board staff

Sponsorship is identified by a footnote on the first page of each report.

INVESTIGATIONS OF THE EFFECT OF TRAVELER ATTITUDES IN A MODEL OF MODE-CHOICE BEHAVIOR

David T. Hartgen and George H. Tanner,
New York State Department of Transportation

This paper summarizes the research conducted by using attitudinal data to predict individual mode choice. In this approach, travel is viewed as a form of human behavior to which the appropriate sociological and psychological principles may be applied. A model of individual mode choice is formulated, based on the traveler's attitudes toward qualitative system characteristics. An index is devised to represent the relative quality of transit and automobile to the traveler and is then related to actual mode choice. The choice model is made operational by merging an attitudinal data set with a file of trips based on the characteristics of trips and tripmakers. A series of tests examines the sensitivity of the model to changes in several qualitative system attributes. Finally, recommendations are made for further studies and refinements that are felt to be necessary if the model is to become an operational planning tool.

•THE USES of attitudinal data in understanding individual travel behavior were investigated during the fall of 1969 by the New York State Department of Transportation. Through a review of the literature on household decision-making, activity patterns, and travel characteristics, a heuristic theory of travel behavior was structured around the social and psychological needs of persons and households. Individual travel decisions (destination, choice of mode, and route) were hypothesized to result from an informal household decision-making process that evaluates the needs of members and assigns to each member certain tasks intended to fulfill those needs. Previously prepared papers (1, 2) describe this approach in more detail.

The research described in this paper concentrated on mode choice within this framework. A model of individual mode choice was devised that included components felt to be essential to understanding travel behavior. These are the characteristics of the traveler and his household, described by socioeconomic variables; the types of activities in which individual household members participate; the distribution of activity sites about the household by alternative modes of travel; and the attitudes of travelers toward the quality of alternative modes. It was felt that using traveler attitudes in the model would permit examination of qualitative (and often subconsciously perceived) factors affecting travel behavior such as comfort, convenience, self-esteem, and personal safety. There is growing evidence (3, 4, 5) that these factors are of considerable importance in mode-choice travel decisions and therefore should be included in choice models. In addition, it was felt that the problems of the demand for new modes, whose attributes may be quite different from those of existing modes, are probably most amenable to solution through consideration of traveler attitudes rather than through extrapolation of engineering measurements.

This paper gives a brief statement of the theory and formulation of a mode-choice model developed to incorporate these approaches and considerations. It also presents a series of tests of the model that is designed to reveal both the sensitivity of the

Sponsored by Committee on Transportation Forecasting and presented at the 50th Annual Meeting.

model's formulation to the qualitative aspects of mode choice and the applicability of attitudinal data to mode choice.

TRAVEL AS A FORM OF HUMAN BEHAVIOR

Each individual has associated with him a set of needs defined by the roles he assumes in his interaction with other persons and groups. Through experience, individuals and groups develop both awareness of and attitudes toward alternate courses of action that may satisfy needs. Through awareness, a person or group recognizes the existence of those particular actions offering some potential for satisfying needs. Attitudes, on the other hand, are the pre-established tendencies for responses toward any of the courses of action identified by awareness. Attitudes and awareness therefore aid individuals by identifying activities and actions that can satisfy needs.

For those activities that require travel for completion, the individual must also consider the characteristics of alternate transportation systems with respect to their usefulness to him. Individual decisions involving travel will depend on this evaluation. In a travel decision, choice of mode, the traveler's view of the transportation systems is particularly important. In studies measuring how persons view transportation modes (3, 4, 6, 7), it is generally agreed that a person's attitude toward new or greatly improved transportation service will be based on his experience with existing service. In other words, he will evaluate new or unfamiliar systems by comparing them with more common ones. This paper suggests a procedure for describing this evaluation. It is hypothesized that the traveler classifies the modes that he perceives to be operating in the transportation system by comparing his attitude toward and awareness of attributes of these modes with his corresponding attitude toward and awareness of attributes of a preconceived ideal (perfect) mode. This allows him to make some statement about the relative quality of alternative modes, both to each other and to the ideal mode. The choice of mode is then determined by evaluating the quality of each mode. Experience from each trip may result in conscious or subconscious changes in the traveler's attitude, thus affecting his subsequent travel decisions.

A mode-choice model based on the previous ideas has been formulated and is being tested. The remainder of this paper deals with the form of this model and some preliminary test results.

MODE-CHOICE MODEL FORMULATION

Recently proposed mode-choice models (8, 9) have suggested that the criteria of traveler mode choice seem to depend on 2 components of perceived system attributes:

1. The importance placed on a given system attribute by a particular traveler for a particular trip (importance expresses awareness); and
2. The degree of satisfaction this traveler has with the ability of each alternative mode to fulfill the requirements of each system attribute (satisfaction expresses attitudes).

The model of traveler mode choice described here is based on the idea that a traveler's attitude toward the modes available for his trip depends on both the importance and the relative quality of a number of aspects of this trip, with each being represented by a number of specific system attributes. The rationale for the model will not be detailed here; the reader is referred to other material (1, 2) for extensive treatment. Put briefly, the model describes a binary choice situation in which the urban traveler chooses between 2 alternative means of travel. The choice is binary because the automobile and transit modes dominate intraurban travel. It is hypothesized that the amount of travel occurring on each mode P_{ik}, for traveler of type k on mode i, is a function of travelers' attitudes toward the quality of alternative modes and the travel times over each network within a particular travel corridor. Travelers' attitudes are measured by an attitudinal index, C_k, whereas the travel times are evaluated by an index of service, SI; therefore, $P_{ik} = f(C_k, SI)$.

The attitude of a traveler toward alternative transportation systems is hypothesized to be a linear combination of his attitude toward each of the factors he perceives to

influence his travel decision. In evaluating the modes available, the traveler considers both the relative importance of each factor and his satisfaction with either mode with regard to that factor. This is expressed as

$$C_k = \sum_q I_{qk} \left(1 - \frac{S_{1qk}}{S_{2qk}}\right) \tag{1}$$

where

C_k = attitudinal index for traveler of type k (Table 3);
I_{qk} = importance of a factor q, defined as the greatest importance placed on any of the i attributes encompassed by factor q, for traveler k— $I_{qk} = \max I_{ik}, i \in q$;
S_{1qk} = satisfaction (generalized benefits) that the traveler experiences with factor q, where S_{1qk} is some function of the i attributes encompassed by factor q toward mode 1 by traveler k; and
S_{2qk} = similar conditions as for S_{1qk}, except that this applies to mode 2.

An index of service was constructed to allow the planner to evaluate major system improvements in specific corridors. The transportation planner, unlike the traveler, is interested in a precise description of transportation system characteristics in terms of engineering measurements. Aspects of the transportation system such as headways and capacities aid the planner in examining system capability, determining capital and operating costs, and relating alternative courses of action to changes in specific system attributes. The service index weighted the ratio of over-the-network travel times for 2 modes by the trip-end density at the nonhome end of the trip.

$$SI = \frac{TP}{AP + ATT} \left[\frac{1}{D_k}\right]^{½}$$

where

SI = service index,
TP = transit door-to-door travel time,
AP = automobile in-vehicle travel time,
ATT = automobile terminal time, and
D_k = density of trip destinations, for a particular trip purpose, at the nonhome end of the trip.

MODE-CHOICE MODEL IMPLEMENTATION

Trips for this study were stratified by trip purpose, automobile availability, household type, and income. These variables were intended to account for the different importances placed on certain system attributes for different types of trips. The levels of the variables are as follows:

Variable	Levels
Trip purpose	Work-school, other
Automobile availability	Automobile available, no automobile available
Household type	Family, nonfamily
Income	\$0-3,999, \$4,000-5,999, \$6,000-9,999, over \$10,000

A discussion of the reasoning behind this structure appears elsewhere (1). Briefly, this combination of variables and levels was felt to be a means of identifying trips by their associate activities. Trip purpose was intended to represent activity purpose, automobile availability was to represent the activity's priority in the household, household type was to represent the idea of hierarchy in household decision-making, and income was to represent the resources of the household.

This trip stratification was then applied to a file of trip data collected by a home-interview survey in Rochester, New York, in 1963. Each trip in the Rochester file was

classified according to one of 32 combinations of the variable levels described previously.

Transportation studies have not collected separate data on the traveler's perception of modal attributes in conjunction with data on travel patterns. One study conducted by the University of Maryland in metropolitan Philadelphia was designed to collect these types of linked data. In that study, the respondent was asked to indicate the level at which each of 2 travel modes satisfied a specific attribute. Independently, the respondent was asked to record the level of importance of this attribute to him. Each set of responses referred to one of 33 attributes that ranged from vehicle cleanliness to travel costs. For a particular trip, both measures of satisfaction and importance were recorded along 7-position Likert scales. This information defined the attitude data file. Table 1 gives a typical list of scores for each of the 33 system attributes for the following combination of variable levels: family, income between $4,000 and $5,999, automobile available, and work trip.

Because of their appropriateness, the Maryland survey data were used as a basic attitude source for the current study. Each response was placed into 1 of the 32 cells that described the trip characteristics by using the levels of the 4 variables described previously. Average scores for each of the 33 attributes were then computed in each of these cells.

The Maryland survey staff found that different combinations of attributes tended to describe different factors important to travel. Reliability is an example of a factor composed of 2 attributes—arrive without accident and avoid stopping for repairs. Although these factors were not entirely independent, it was felt that for modeling purposes they could serve to differentiate system attributes. Each attribute investigated by the Maryland survey could be similarly described by a number of transportation system variables.

Table 1 gives the combination of the 33 system attributes into 11 factors for a particular type of trip. Those attributes representing factors were then used to calculate attitudinal indexes according to Eq. 1. Table 2 gives an example of this calculation using the same combination of variable levels as used for Table 1. The entire set of indexes for all 32 cells is given in Table 3.

The attitudinal data devised in this manner were then applied to the analysis of modal-split estimates in Rochester. Because the attitudinal data had been obtained from individuals living in the Philadelphia metropolitan area, it was necessary to relate the observed values from the Philadelphia survey to data from similar trips recorded in the Rochester survey. This was accomplished by assigning the indexes given in Table 3 to trips of a similar type made in the Rochester area. Thus, each trip recorded in the Rochester survey was assigned one of the attitudinal indexes given in Table 3 based on trip and household characteristics.

The procedure for developing estimates of mode use was based on the concept of response surfaces used in several transportation studies (10). These surfaces can be described as 3-dimensional diversion curves. In general, a response surface is constructed by arraying the percentage of trips by a mode, usually transit, against several demographic, geographic, or system variables, or against all three. Four response surfaces were created for combinations of 2 trip purposes with 2 automobile-availability categories. On each surface the percentage of transit trips was arrayed by the service index and the attitude index as shown in Figure 1.

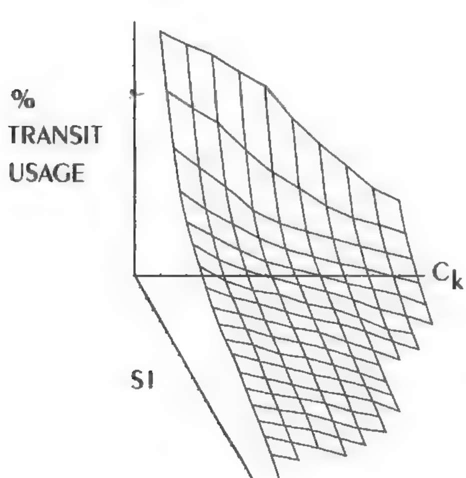


Figure 1. Typical response surface.

TABLE 1

SATISFACTION AND IMPORTANCE SCORES FOR SYSTEM ATTRIBUTES BY FACTOR

Factor	System Attribute		Perceived Satisfaction[a]		Perceived Importance[a]
	Number	Description	Transit	Automobile	
Reliability	28	Arrive without accident	6.0	6.1	6.2[b]
	33	Avoid stopping for repairs	5.2	5.5	4.9
Travel time	1	Arrive in shortest time possible	2.5	5.8	5.9
	7	Travel in light traffic	3.9	4.8	4.1
	11	Arrive at intended time	4.4	5.9	6.4[b]
	12	Arrive in shortest distance	4.3	6.4	6.1
	17	Avoid changing vehicle	4.3	6.2	5.0
	18	Ride in safest possible vehicle	5.6	5.9	6.1
	24	Travel as fast as possible	4.3	5.6	6.4
Weather	5	Have protection from weather while waiting	3.0	5.2	5.0
	2	Vehicle unaffected by weather	4.6	6.0	5.6[b]
Cost	3	Total trip cost	4.5	5.7	5.0[b]
	13	One-way cost of 25 cents instead of 35 cents	4.9	5.5	4.2
	22	One-way cost of 25 cents instead of 50 cents	4.6	5.7	4.0
	29	Per-mile cost of 3 cents instead of 15 cents	4.9	5.6	4.5
Vehicle condition	10	Ride in clean vehicle	4.8	5.4	5.6[b]
	20	Ride in new modern vehicle	5.1	5.8	5.1
Personal safety	30	Avoid unfamiliar area	5.0	5.7	4.3[b]
Self-esteem	6	Ride in uncrowded vehicle	4.0	6.0	3.8
	14	Have feeling of independence	4.0	6.2	5.1
	26	Avoid waiting more than 5 min	4.0	6.3	5.8[b]
	27	Ride comfortably	4.8	5.9	5.5
	31	Have pride in vehicle	4.9	4.8	3.6
	32	Avoid riding with strangers	5.0	5.7	4.2
Diversions	4	Listen to radio	5.0	5.5	4.8[b]
	8	Take along family and friends	4.0	6.1	3.9
	9	Ride with people who chat	4.7	5.6	4.0
	15	Look at scenery	4.5	5.6	3.9
	21	Ride with friendly people	4.8	6.4	4.7
	23	Ride with people you like	4.6	6.4	4.3
Convenience	16	Avoid walking more than a block	3.9	6.1	5.5[b]
Packaging	19	Have package and baggage space	4.4	6.2	4.3[b]
Fare payment	25	Need not pay fare daily	5.3	6.1	3.7[b]

Source: University of Maryland Study.

[a]Variable levels: work trip, automobile available, family, income \$4,000-\$5,999.

[b]Maximum importance within each factor, q.

TABLE 2

EXAMPLE OF ATTITUDINAL INDEX CALCULATION

Factor q	System Attribute	Satisfaction: Transit S_{1qk}	Satisfaction: Automobile S_{2qk}	$1 - \frac{S_{1qk}}{S_{2qk}}$	Importance I_{qk}	Added Value $I_{qk}\left(1 - \frac{S_{1qk}}{S_{2qk}}\right)$
1	Arrive without accident	6.0	6.1	0.016	6.2	0.10
2	Arrive at intended time	4.4	5.9	0.254	6.4	1.63
3	Vehicle unaffected by weather	4.6	6.0	0.234	5.6	1.31
4	Total trip cost	4.5	5.7	0.211	5.0	1.06
5	Ride in clean vehicle	4.8	5.4	0.111	5.6	0.62
6	Avoid unfamiliar area	5.0	5.7	0.123	4.3	0.53
7	Avoid waiting more than 5 min	4.0	6.3	0.365	5.8	2.11
8	Listen to radio	5.0	5.5	0.091	4.8	0.44
9	Avoid walking more than a block	3.9	6.1	0.361	5.5	1.98
10	Have package and baggage space	4.4	6.2	0.290	4.3	1.25
11	Need not pay fare daily	5.3	6.1	0.132	3.7	0.49
Total						11.51[a]

Note: Variable levels are work trip, automobile available, family, income $4,000 to $5,999.

[a]Attitudinal index.

Basically, the model operates by relating a change in some specific system attribute (e.g., vehicle cleanliness) to a change in traveler satisfaction with that attribute. This change in satisfaction may result in a change in the traveler's attitude toward the system, which is measured by the attitudinal index. When applied to the response

TABLE 3

ATTITUDINAL INDEXES BY TRIP CLASS

Household Type	Income (dollars)	Trip Characteristics: Automobile Available: Work	Automobile Available: Nonwork	No Automobile Available: Work	No Automobile Available: Nonwork
Non-family (single persons, roommates)	0-3,999	8.15	14.12	12.14	8.91
	4,000-5,999	7.22	15.76	12.97	9.75
	6,000-9,999	5.98	14.29	19.17	10.74
	10,000+	6.81	18.23	19.77	12.59
Family	0-3,999	13.67	13.95	2.80	3.96
	4,000-5,999	11.51	12.62	6.23	5.55
	6,000-9,999	9.54	12.44	8.41	5.93
	10,000+	9.86	15.38	12.22	9.76

Source: University of Maryland study.

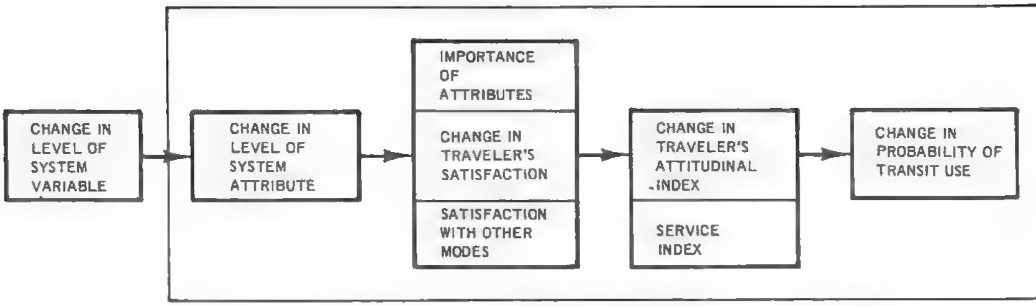


Figure 2. Schematic operation of the model.

surface, this change in the attitudinal index is interpreted as a shift in the trip's position on the surface, resulting in a change in the probability that this trip will be made by transit. When this probability is applied to a large number of trips, the result will be a change in the number of transit users. This process is shown in Figure 2.

The key step in this sequence (Fig. 2) is the description of the relationship between traveler attitude toward a system attribute and the value of specific variables used to describe the attribute. In theory, this may be accomplished by relating specific levels of system variables to attitudinal scores.

There are a number of problems associated with this approach, but foremost is the absence of data that can aid in specifying the relationship between a system attribute and its various variable measures. Therefore, we were forced in the following tests to assume that changes occurred directly in the satisfaction levels of attributes and to use these as the means of inducing attitude change in the model.

A second assumption is inherent in the test implementations of the model. It concerns the acceptability of using attitude data from the Philadelphia survey merged by trip and household characteristics with travel data from the Rochester survey. This second assumption, however, is made in light of some evidence from urban areas that suggests that attitudes are influenced primarily by families and groups rather than by geographic location (11).

Because of the nature of these 2 assumptions, test results are not subjected to any rigorous statistical analysis. Rather, output from the following tests is intended to illustrate the potential of this model form.

DESCRIPTION OF TESTS AND RESULTS

The model as described was applied first to reproducing estimates of mode use as determined by the Rochester survey travel files. Table 4 gives values for transit trips by ring of trip-maker's residence and by ring of trip origin. For purposes of analysis, study areas are described by a series of concentric rings emanating from the CBD. Trips by ring of trip-maker's residence are defined as transit trips that have a trip purpose of either the home as origin or the home as destination in the identified ring. Transit trips by ring of trip origin, on the other hand, are simply those transit trips that originate for any purpose in the indicated ring. Examination of the data given in Table 4 reveals the relative contributions of each of the rings to transit ridership in Rochester as estimated by the model.

Results from the tests are given in Table 5. Table 5 gives the response to hypothetical differences in terms of differences in transit trips by ring of trip-maker's residence and the response by ring of trip origin. These 2 views of ridership response show which residence rings are most affected by the test changes and which ones are affected by travel patterns of the residents.

Test 1

The first test investigates the effect within the model of an assumed increase in the cleanliness of buses (attribute 10, Table 1), as interpreted by an increase in the

TABLE 4

SUMMARY OF MODEL-ESTIMATED TRANSIT TRIPS BY RING OF TRIP-MAKER'S RESIDENCE AND TRIP ORIGIN

Ring	Trips by Ring of Trip-Maker's Residence		Trips by Ring of Trip Origin	
	Number	Percent	Number	Percent
0	2,376	9.2	16,051	17.1
1	4,729	10.4	10,888	11.5
2	30,194	10.8	24,842	8.3
3	28,265	8.3	26,590	6.2
4	18,248	7.3	15,246	5.2
5	5,075	5.1	3,083	2.6
6	3,403	3.5	1,857	1.6
7	1,672	5.1	784	2.9
Total	93,962	8.3	100,358	7.0

satisfaction level of travelers with the cleanliness of the transit vehicle. For this test, the changes given in Table 6 of the percentage of persons satisfied with transit vehicle cleanliness were assumed. (As noted, the figures given in Table 6 result from applying Philadelphia attitudes to Rochester trip data.)

These improvements were assumed to apply to all travelers in the urban area, because bus cleanliness was assumed to be uniform throughout the study area. For each level of system-wide transit vehicle cleanliness, higher income groups express much less satisfaction than lower income groups do. The changes made to the percentage of persons satisfied were based on the assumption that those groups who express the greatest satisfaction with present vehicle cleanliness would be less affected by cleaner vehicles than those groups who registered less satisfaction with this variable.

The percentage of changes in satisfaction resulting from the improvements in vehicle cleanliness were then used in calculating a new set of attitudinal indexes. When the new attitudinal indexes were applied to the response surfaces discussed previously, corresponding increases in transit use resulted. The magnitude and location of the transit use increases are given in Table 5 (Test 1). Most of the resulting increase in transit use developed in this test occurred in rings 2, 3, and 4 because of the high transit potential of the area bordering the CBD. These rings seem to possess the greatest number of 1-car households. On the other hand, the greatest positive percentage differences in transit use are in rings 5, 6, and 7. These outer rings are essentially suburban areas having a substantial number of higher income households that, through the test conditions input to the model, had experienced the greatest increase in satisfaction level. By contrast, the CBD and ring 1, which contain a large number of lower income households, exhibit the lowest response to clean buses, when viewed in terms of percentage difference to initial conditions.

Although arbitrary, the magnitude of these results appears reasonable. One would expect only a small increase in transit use to result from implementing a relatively unimportant attribute such as vehicle cleanliness. Nevertheless, the model does seem capable of evaluating the effect of qualitative components not normally included in other modal-split mechanisms.

Test 2

The effect of a downtown transportation terminal was investigated in a second test. It is felt by some analysts that such a terminal, if well designed, could increase the use of a transit system by providing more efficient service of better quality. In this test, the terminal was evaluated in terms of its effect on system attributes.

TABLE 5

RING SUMMARY OF SIX TESTS

Ring	Calibrated Base	Test 1 Clean Buses		Test 2 Downtown Terminal		Test 3 No Fare		Test 4 New Vehicles		Test 5 Equal Access		Test 6 Equal Satisfaction	
	Volume	Volume	Difference (percent)	Volume	Difference (percent)	Volume	Difference (percent)	Volume	Difference (percent)	Volume	Difference (percent)	Volume	Difference (percent)
Transit Trips by Ring of Trip-Maker's Residence													
0	2,376	2,453	3.2	2,515	5.9	2,667	12.3	3,007	26.7	3,587	51.0	3,720	56.5
1	4,729	4,864	2.9	4,764	0.7	5,170	9.3	5,629	19.0	6,995	48.0	7,049	49.0
2	30,194	31,577	4.6	30,688	1.6	32,993	9.3	36,959	22.3	44,034	46.0	45,738	51.5
3	28,265	29,865	5.7	28,772	1.8	31,272	10.6	36,116	27.8	45,046	59.5	45,578	61.5
4	18,264	19,515	6.9	18,493	1.3	20,320	11.4	23,908	31.0	27,728	52.0	30,128	65.0
5	5,075	5,449	7.4	5,155	1.6	5,681	12.0	6,721	32.5	9,833	94.0	8,679	71.0
6	3,403	3,639	6.9	3,473	2.1	3,782	11.1	4,401	29.1	7,375	117.0	5,657	66.0
7	1,672	1,807	8.1	1,714	2.5	1,875	12.1	2,166	29.5	4,031	141.0	2,834	69.5
Total	93,962	99,169	5.6	95,974	1.7	103,760	10.4	118,906	26.6	148,629	58.2	149,383	58.0
Transit Trips by Ring of Trip Origin													
0	16,051	16,482	2.7	16,890	5.2	17,254	7.5	18,806	17.2	17,939	11.8	21,634	34.8
1	10,888	11,311	3.9	10,927	0.3	11,838	8.7	13,204	21.3	15,490	29.8	16,255	49.4
2	24,842	26,120	5.1	25,118	1.1	27,324	10.0	30,938	24.5	41,864	68.5	38,653	55.6
3	26,590	28,220	6.1	26,863	1.0	29,488	10.9	34,089	28.2	48,857	83.8	43,514	62.9
4	15,246	16,373	7.4	15,369	0.8	17,000	11.5	19,974	31.0	24,960	60.1	25,131	64.8
5	3,083	3,335	8.2	3,119	1.2	3,479	12.8	4,157	34.8	7,776	152.0	5,442	76.5
6	1,857	1,988	7.1	1,891	1.8	2,045	10.1	2,394	29.0	4,964	167.0	3,109	67.5
7	784	848	8.2	804	2.5	878	12.0	1,008	28.6	1,877	140.0	1,329	69.5
Total	100,358	105,565	5.2	101,972	1.6	110,153	9.7	125,299	24.8	164,310	63.7	155,777	55.2

TABLE 6

CHANGES ASSUMED FOR TEST 1

Income Level (dollar)	Persons Initially Satisfied (percent)	Persons Newly Satisfied (percent)	Change (percent)
0-3,999	45	50	+11
4,000-5,999	40	46	+12
6,000-9,999	30	40	+33
Over 10,000	25	38	+48

Transit satisfaction levels of certain attributes selected from the list given in Table 1 were assumed to change, as given in Table 7. The justification for selecting these items is as follows: The downtown terminal was assumed to be strategically placed so as to shorten the travel time for the greatest number of people traveling to the CBD. A terminal facility would certainly provide protection from the weather through its structure and possibly through its location, which could eliminate the need for outdoor transfers to reach intended destinations within the CBD. More ridership to the CBD and, because service is kept constant in this test, more crowding of vehicles may be anticipated. Knowing the social safety afforded by a downtown terminal facility, the traveler would be more willing to bring along his family and friends. These ideas are reflected in the changes in the satisfaction levels of the various income groups as given in Table 7.

The new satisfaction levels created as a result of the test changes were then used to develop new attitudinal indexes that led to the resulting increases in transit use as given in Table 5 (Test 2). Examination of both trips by ring of trip-maker's residence and by ring of trip origin reveals that the initial CBD-origin trips increased almost 6 percent. A modest increase occurred in all rings, except in ring 1 that surrounds the CBD. An increase in CBD transit use was expected because CBD travel would be stimulated by the terminal. The lower response found in ring 1 may be due to the penalizing effects of a downtown terminal in increasing walking distances for riders between ring 1 and the CBD. In terms of absolute ridership, increases are most apparent in rings 2, 3, and 4 where the greatest ridership potential exists.

TABLE 7

CHANGES ASSUMED FOR TEST 2

System Attribute		Change in Satisfaction Level by Income Level, percent			
Number	Description	\$0-3,999	\$4-5,999	\$6-9,999	\$10,000+
1	Arrive in shortest time possible	5	5	10	10
5	Have protection from weather while waiting	40	50	50	60
6	Ride in uncrowded vehicle	- 5	-10	-10	-10
8	Take along family and friends	10	10	10	10
11	Arrive at intended time	10	15	20	20
16	Avoid walking more than a block	-10	-10	-20	-20
17	Avoid changing vehicle	2	5	5	5
20	Ride in new modern vehicle	10	20	25	30
25	Need not pay fare daily	10	10	5	5
26	Avoid waiting more than 5 min	10	15	20	20
30	Avoid unfamiliar area	20	20	25	30

Test 3

The third test investigates the influence of a no-fare or free transit system on transit ridership. Suppose, it is suggested, that transit fares are reduced while the level of service, in terms of routes, headways, and vehicles, remains constant. In this test the satisfaction of transit riders with fares was maximized. It was assumed that everyone would be highly satisfied with a zero-cost fare. The transit satisfaction level of all riders with all items in the trip cost factor (Table 1) was set to 7.0. A new set of attitudinal indexes was computed and resulted in the increases in transit use as given in Table 5 (Test 3). Overall, transit use increased approximately 10 percent. Comparison of the percentage difference by ring trip-maker's residence and by ring of trip origin shows the CBD resident to be more affected by the omission of a fare than the CBD traveler. Although this condition exists in some other rings, it is to be most expected in the CBD and surrounding area. It is apparent that a fare charge consumes a greater proportion of a lower income salary than a higher income salary; hence, a substantial fare reduction could be expected to have the greatest impact on ridership from the lower income class. In the more suburban rings, the incentive to use transit brought about by the fare reduction is not large enough to influence markedly transit ridership. The outer rings have inferior transit service, and a decrease in fare would do little to reduce the inconvenience of transit riding present in these areas.

Test 4

The fourth test examines the effect of substantial improvement in the vehicle condition of the transit system vehicles. Such an improvement could be made through the purchase and maintenance of all-new, noise-free, pleasant-smelling, clean, highly reliable buses. If such a fleet could be purchased and put into operation at existing service levels, thereby replacing all existing vehicles, some increase in patronage would be expected. These substantial improvements to vehicle conditions would probably be widely recognized, especially if these improvements were combined with an extensive public relations and advertising campaign. The result of this effort would be to make almost all persons very satisfied with the transit vehicle condition.

It was assumed that the test changes could be implemented through the changes to the individuals' satisfactions given in Table 8. We assumed that all riders are completely satisfied with the newness and cleanliness of the vehicles. New vehicles would probably also possess other attributes conducive to ridership. Seats and spacing would be improved over present systems. Confidence in safety and reliability could be expected to increase. Pride in the vehicle could be expected to increase. Finally, because the trend in vehicle design appears to be directed toward the inclusion of larger transparent areas, some increase in the satisfaction expressed by the transit rider with his ability to look at the scenery might be expected.

TABLE 8

CHANGES ASSUMED FOR TEST 4

System Attribute		Change (percent)
Number	Description	
10	Ride in clean vehicle	700[a]
20	Ride in new modern vehicle	700[a]
27	Ride comfortably	60
28	Arrive without accident	50
33	Avoid stopping for repairs	50
31	Have pride in vehicle	40
15	Look at scenery	10

[a]Maximum condition

These changes to the transit satisfaction were applied across all income levels and used to compute a new set of attitudinal indexes. The results of using the test information in the model are given in Table 5 (Test 4). The overall increase in ridership is approximately 25 percent. The CBD and rings 5, 6, and 7 are most influenced by new vehicles. The percentage increase in the outer rings (5, 6, and 7) is above the study area average. The absolute increase in ridership that can be attributed to new vehicles is much lower in these rings than in rings 2, 3, and 4. New vehicles along existing routes with existing system characteristics do little, if anything, to affect accessibility; hence, there is a lower absolute response in the suburban rings (5, 6, and 7).

Test 5

The fifth test is an attempt to view the impact of attitudes toward travel modes on existing transit patronage. This test is accomplished by neutralizing the effect of the travel-time ratio in the service index—setting it equal to unity and creating a condition of equal accessibility (equal time). Automobile availability, trip-end density, and attitude index become the discriminants of transit use when the travel-time ratio equals one. If the 2 modes were perceived equally, the expected modal split would be 50 percent transit and 50 percent automobile. For this test, the resulting modal split can be expected to deviate by some fixed amount from the theoretical 50-50 value, with transit being substantially lower than the automobile (possibly 15-85). The lower use for transit would result from the fact that most people expressed attitudes suggesting that they were more satisfied with the automobile than with public transit for reasons other than accessibility.

This hypothesis was tested by equating system attributes on the service index and observing the model's estimate of the resulting transit use. The results are given in Table 5 (Test 5). As might be anticipated, those rings adjacent to the CBD, predominantly composed of households with 1 automobile, exhibit substantial increases in the number of transit riders. The suburban rings, with higher incomes, more automobiles available, and more pronounced dissatisfaction with transit service, register more modest ridership increases. The overall increased transit use is nearly 60 percent.

Equal access appears to have more of an effect on trips by ring of trip origin than on trips by resident dwelling ring. This is apparent for rings 2 through 6 where there is a marked difference between the 2 respective percentage difference values.

Test 6

For the fifth test, the service index was neutralized. The sixth test investigates the influence of travel time on mode use through the neutralization of the attitude index, creating conditions of equal satisfaction. Equivalency is established by setting the satisfaction of a factor for transit equal to the satisfaction of that same factor for automobile. In this manner each traveler is assumed to have equally favorable attitudes toward both modes. The conditions of equal satisfactions, when summed over all factors in the computation of the attitude index (Eq. 1), result in a value of zero. Structured in this manner, the test studies the effect on transit use of these variables: travel-time ratio, trip-end density, and automobile availability. In this case, the model operated as a diversion curve, based on the respondent's values for time and money.

The results of this test are given in Table 5 (Test 6). In the suburban rings (5, 6, and 7), attitudes appear to have had a more substantial influence on determining transit use than in the inner rings because the greatest percentage increase occurs in this outer area. The overall increase from the total initial ridership is nearly 55 percent. Although the model is used as a diversion curve, this test is not typical of the more commonly accepted diversion curves. The usual diversion curve developed without distinct consideration of qualitative factors had these factors implicit in its construction. The curve used in this test, on the other hand, is devoid of qualitative evaluations and is hinged on cost, trip-end density, and automobile availability. When the components of the service index in this case are examined, the fare cost of transit is less of an influence on mode choice of suburban residents than the time cost of transit travel, which is usually prohibitive.

In all of these tests, trip-end density, although noted, has not been a controlling variable because of a constant trip distribution for all 6 tests.

SUMMARY AND DISCUSSION

The model develops an estimate of expected transit patronage—the proportion of zone-to-zone trips that will use the transit mode—as a function of 4 general factors:

1. Operating characteristics of both transit and automobile systems (such as speed, parking charges, headway, and trip density) as indicated by the service index;
2. Stratification of trips (purpose, automobile availability);
3. Demographic aspects of the region (spatial distribution of households by income, household structure); and
4. Attitudes of travelers toward abstracted system attributes as indicated by the attitude index.

The model tests described were intended to illustrate the potential uses for this tool and to provide insight into the rationality of model response to specific conditions.

In the first 4 tests of the model, emphasis was placed on the effect of attitudes on mode-choice travel behavior. These initial tests were concerned with the impact of various qualitative improvements to the transit system. The tests examined the effect of clean buses, a downtown terminal, no fares, and new vehicles, all of which resulted in increases in transit use. The greatest increases in patronage resulted from the no-fare and new-vehicle tests.

The last 2 tests were intended to operate the model near its tolerable limits. These tests examined the influence of equal access and equal satisfaction. Increased transit mobility, or equal access (Test 5) appeared as a stronger patronage stimulant than fare reduction (Test 3). This observation is in agreement with the result of at least 1 public transportation demonstration project (12). Creating more favorable attitudes toward transit (Test 6) appears to be nearly as important as increasing transit accessibility (Test 5). This observation is based on a comparison of the percentage differences of Tests 5 and 6 (Table 5).

Through incorporation of the traveler's attitude toward the transportation system, the planner may "see" the systems from the traveler's viewpoint, and (theoretically) relate these attitudes to specific quantifiable physical variables that are of concern to the system designer.

A model of the type presented in this paper is also capable of estimating the patronage that may be attributed to new modes by extending present attitudes toward abstracted features of existing modes and projecting them to their new mode counterpart, which is much in the same manner that a person relates past experience to his analysis of the future.

Successful application of the model to this problem would, of course, require knowing the relation between satisfaction levels and levels of specific system variables describing new modes. As in most mode-choice models, the formulation and tests described represent many compromises and are therefore somewhat less than ideal. The data on which the model is based only allow operation of the model as a valuable but limited research tool. It is not possible, at the moment, to apply the model with confidence to actual planning situations. With these comments in mind, one may identify the shortcomings of the model with 2 major areas—model formulation and available data.

Examination of all the assumptions made concerning the formulation of the model would be appropriate. A few key areas that should be examined in greater detail are the relation of travel to activities and household needs, which is the mechanism that governs travel decisions, and the travelers' perception of transportation system attributes.

Similarly, there are many pieces of additional data that should be gathered. Improved data would set the stage for examining the sensitivity of the model to changes in input parameters. On the surface, the model appears to react in a credible manner, but considerably more testing must be done before the model is applied to the problems of a particular city. Two immediate requirements are to gather engineering and

attitude data together in various cities to gain knowledge of the relationship between system operations and traveler attitudes and expand the 33 specific items to permit the consideration of a greater number of problems. One can hope that the result would be improved understanding of travel as a behavioral phenomenon.

ACKNOWLEDGMENTS

We appreciate the assistance of Allen N. Nash of the University of Maryland in obtaining the attitudinal data for this study. We also thank John K. Lussi and G. Frederick Young for helpful comments throughout the project, Ellen Lindop for editorial assistance, Anne M. Edmonds and Pamela Houghtaling for manuscript preparation and technical assistance, and Estelle L. Tweedie for programming and operation of the model.

REFERENCES

1. Hartgen, D. T., and Tanner, G. H. Individual Attitudes and Family Activities: A Behavioral Model of Mode Choice. High Speed Ground Transportation Jour., Vol. 4, No. 2, Sept. 1970.
2. Hartgen, D. T., and Tanner, G. H. Behavioral Model of Mode Choice. New York State Department of Transportation, Preliminary Res. Rept. 19, March 9, 1970.
3. Brunner, G. A., et al. User Determined Attributes of Ideal Transportation Systems: An Empirical Study. Dept. of Business Administration, Univ. of Maryland, College Park, June 1966.
4. Paine, F., et al. Consumer Conceived Attributes of Transportation: An Attitude Study. Dept. of Business Administration, Univ. of Maryland, College Park, June 1967.
5. Transit Usage Forecasting Techniques: A Review and New Directions. CONSAD Research Corp., Pittsburgh, Penn., April 1968.
6. McMillan, R. K., and Assael, H. National Survey of Transportation Attitudes and Behavior: Phase 1, Summary Report. NCHRP Rept. 49, 1968.
7. McMillan, R. K., and Assael, H. National Survey of Transportation Attitudes and Behavior: Phase 2, Analysis Report. NCHRP Rept. 82, 1969.
8. Sommers, A. N. Towards a Theory of Traveler Mode Choice. High Speed Ground Transportation Jour., Vol. 4, No. 1, Jan. 1970, pp. 1-8.
9. Golob, T. F. The Survey of User Choice of Alternate Transportation Modes. High Speed Ground Transportation Jour., Vol. 4, No. 1, Jan. 1970.
10. Fertal, M. J., et al. Modal Split: Documentation of Nine Methods of Estimating Transit Usage. U.S. Department of Commerce, Bureau of Public Roads, 1966.
11. Trezpacz, M. Household Trip Production: A Comparison of Three Urban Areas in Upstate New York. Rensselaer Polytechnic Institute, Troy, New York, unpublished PhD dissertation, Dec. 1969.
12. Pignataro, L. J. Summary of Results From HHFA/HUD Mass Transit Demonstration Studies. In Urban Mass Transit Planning, (Hamburger, W. S., ed.), ITTE, Univ. of California, Berkeley, 1967, p. 71.

DIVERSION MODEL FOR ESTIMATING HIGH-SPEED RAIL USE

John F. DiRenzo and Louis P. Rossi,
New York State Department of Transportation

One proposed solution for many intercity transportation problems affecting densely populated travel corridors in the nation is the provision of high-speed rail service. An important parameter that must be estimated to determine the feasibility of this proposed service is anticipated use. This paper presents a diversion model that estimates high-speed rail passenger use and is applicable to studies attempting to identify corridors that would likely support such service and to determine whether previously proposed high-speed rail service offers sufficient potential to justify conducting detailed feasibility studies. The potential market for high-speed rail service is estimated first by stratifying intercity trips on each nonrail mode by termination point (CBD or non-CBD), by trip purpose (business or nonbusiness), and by group size for automobile trips. The characteristics of the resulting market segments are then analyzed by using travel time and cost data for high-speed rail and each competing mode to determine whether each market segment is completely divertible, possibly divertible, or nondivertible to high-speed rail service. The diversion model provided satisfactory results in a concept study of high-speed rail service in 3 travel corridors within New York State. Certain benefits of this technique are that it is simple to understand and apply, it does not have to be calibrated in the traditional manner, and it can be applied by using available travel data.

•THE PROVISION of high-speed rail passenger service has been suggested as a possible solution for many intercity transportation problems in certain densely populated corridors in the nation. Widespread consideration is being given to this transportation concept. The U.S. Department of Transportation and the Penn-Central Railroad, cooperating in the Northeast Corridor Transportation Project, are attempting to evaluate the feasibility and impacts of providing high-speed rail service between New York City and Boston, and between New York City and Washington. Currently, New York State is also conducting a study of the feasibility of high-speed rail service between New York City and Buffalo (via Albany). Similar studies are being sponsored by the Commonwealth of Pennsylvania and the Southeastern Pennsylvania Transportation Authority for proposed service between Philadelphia and Harrisburg, and by the New England Regional Commission for proposed service between Boston and New York City. Furthermore, legislation has been passed by Congress establishing a quasi-public corporation to provide or ensure the continuation of high-quality, perhaps high-speed, intercity rail passenger service in certain heavily traveled corridors.

High-speed rail service is often recommended because it can provide direct service between downtown areas of the cities served. Also, because high-speed rail service could probably utilize existing rail rights-of-way, new and costly rights-of-way would not have to be constructed, and transportation-related air pollution could likely be reduced. Because travelers would be diverted from other modes, congestion on other

---

Sponsored by Committee on Transportation Forecasting and presented at the 50th Annual Meeting.

modes and facilities would be lessened, and, ideally, a more balanced transportation system would be created.

It is likely that many agencies will conduct concept or preliminary feasibility studies to evaluate the applicability of high-speed rail service within their particular corridors. Anticipated use is one of the basic parameters that these studies must estimate so that feasibility or, at least, the promise, of high-speed rail service can be determined. The purpose of this paper is to present a diversion model for estimating high-speed rail use that can be used in studies that attempt to identify corridors that would likely support such service and in studies that attempt to verify that proposed high-speed rail service in a particular corridor demonstrates sufficient potential to justify additional expenditures on detailed feasibility studies. This model was developed for use in a concept study of high-speed rail service in New York State (1, 2).

Use estimates are a necessary input for estimating operating revenues and costs, user benefits and costs, and other significant factors affecting the feasibility of new or improved rail service. When combined with an evaluation of equipment alternatives, operational problems, needs for physical facilities, and alternative means of financing and implementing the service, this information provides a sound basis for determining the feasibility of high-speed rail service in heavily traveled corridors.

DIVERSION MODEL TO ESTIMATE HIGH-SPEED RAIL USE

A diversion model was used to estimate high-speed rail use in a concept study of such service in New York State. Five factors were explicitly considered in the model: travel time, travel cost, trip purpose, destination location, and, for automobile trips, group size of intercity trips. The formulation of the diversion model and the data used to apply it are given in this section. The objective of the study in which the model was used is also discussed. Estimated high-speed rail use for 1968 will be discussed later for 3 travel corridors in New York State.

The objective of the New York State study was to determine whether the high-speed rail service concept merited further study and, if so, the appropriate geographic extent of subsequent detailed market and engineering feasibility studies, leading to the possible implementation of such service. This determination was based on order of magnitude estimates of ridership for 1968 and on the corresponding direct and indirect benefits and costs of implementing high-speed rail service in 3 key overlapping travel corridors of New York State (Fig. 1): (a) New York City to Buffalo via Albany, (b) New York City to Montreal via Albany, and (c) New York City to Albany.

Diversion Model

Potential high-speed rail ridership in the 3 New York State corridors was estimated in 2 steps. First, the total number of intercity trips made on each nonrail mode between each city pair in the corridors was stratified into market segments by using trip characteristics identified in other studies as important determinants of modal split (3, 4, 5, 6, 7). This was necessary because the attractiveness of high-speed rail service is likely to vary among different segments of the intercity travel market. Second, the market segments for each mode were analyzed in conjunction with travel time and cost data for high-speed rail and each competing mode to determine which segments are likely to divert to high-speed rail service. This determination was made by using a set of reasoned decision rules sensitive to the previously mentioned determinants of modal split.

The 3 factors used to segment the intercity trips on each nonrail mode are shown in Figure 2. Because of the city-center nature of the proposed rail service, it is important to separate trips destined to the CBD from those destined to non-CBD locations. The distinction is also made between business and nonbusiness trips because business and nonbusiness travelers generally value their time and the cost of a trip differently. Furthermore, when estimates are made of diversions from automobile to high-speed rail service, it is important to distinguish between single- and multi-person trips. Multi-person trips are less likely to divert from automobile because of the additional cost incurred when two or more persons use public transportation.

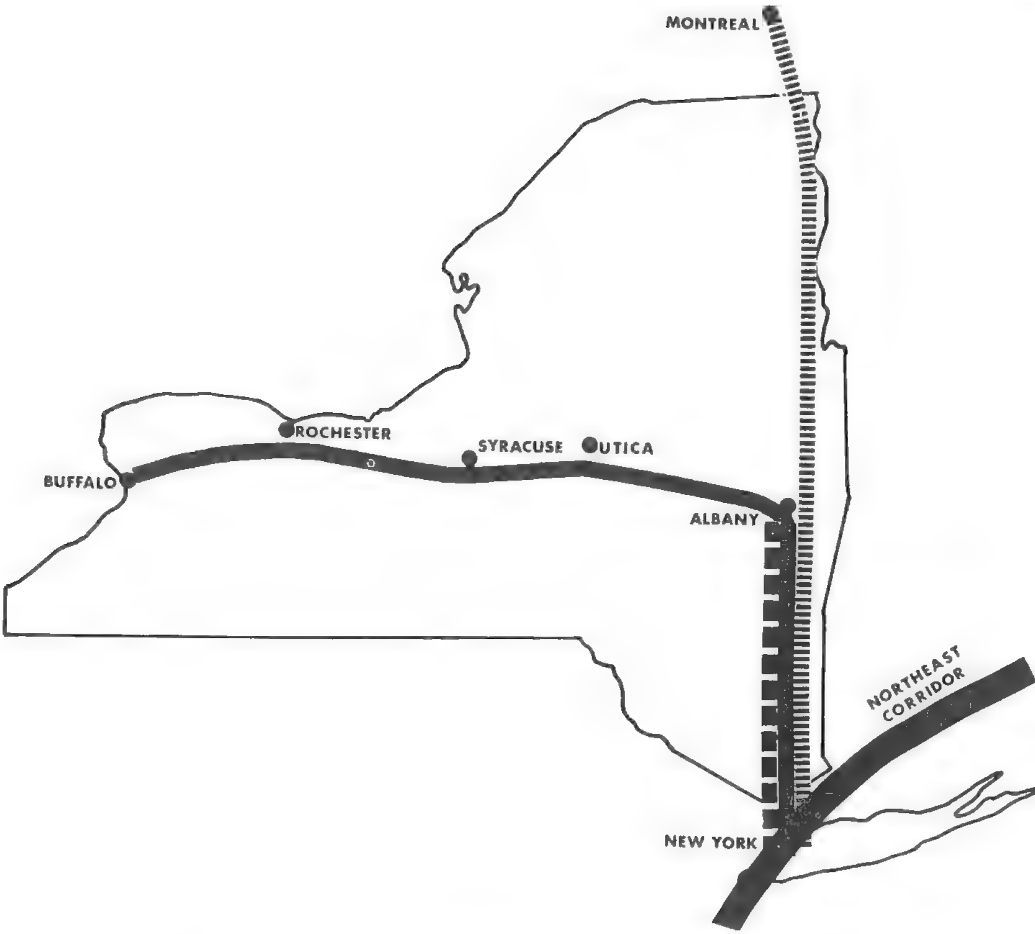


Figure 1. Alternative high-speed rail corridors in New York State.

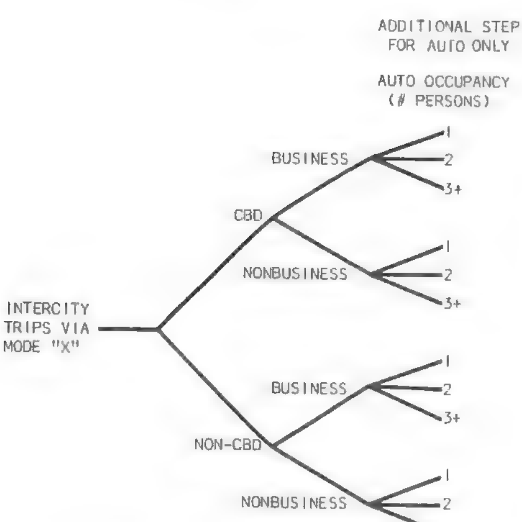


Figure 2. Procedure used to stratify intercity trips.

For each city pair, each identified market segment was classified into 1 of 3 diversion categories by applying the following set of reasoned decision rules:

1. A market segment is considered to be completely divertible to high-speed rail service if the train is faster and less expensive than the mode used at that time.
2. A market segment is considered to be nondivertible to high-speed rail service if the train is slower and more expensive than the mode used at that time.
3. A market segment is considered to be possibly divertible to high-speed rail service (a) if the train is faster and more expensive than the mode used at that time or (b) if the train is slower and less expensive than the mode used at that time.

TABLE 1

DATA USED TO DEVELOP AND APPLY THE DIVERSION MODEL

Item	Automobile	Bus	Air	Rail
Travel volume (person trips)	x	x	x	x
CBD and non-CBD split	x		x	
Trip purpose	x	x	x	
Car occupancy	x			
Door-to-door travel time				
Main line	x	x	x	x
Terminal		x	x	x
Access		x	x	x
Fare	x[a]	x	x	x
Service frequency		x	x	x

[a] Automobile operating cost.

The reasoned decision rules form the basis of the diversion model.

To develop and later apply the diversion model required that data be obtained describing the service provided by each existing intercity mode serving the corridors and the nature and number of trips made between each city pair on each mode. These data were collected from a wide variety of secondary sources; no new surveys were conducted for this study. Air-travel data were obtained from carrier ticket records, passenger surveys, and schedules. Automobile-travel data were obtained from files of earlier urban transportation surveys in the major cities of the state and from other travel data records maintained by the Department of Transportation. Intercity bus and rail companies provided travel-volume data for their respective modes. In some instances, travel data for the base year (1968) were not available, making it necessary to factor the next most recent information to the base year. Several travel volumes between smaller upstate cities were estimated to complete the city-pair trip matrices needed for the analysis of potential ridership.

Table 1 gives a summary of the items of information collected for this study. Travel volumes, in terms of person trips, between each city pair were gathered for each mode. Information was also obtained on the CBD and non-CBD orientation of air and automobile trips, trip purpose of nonrail travelers, and, for automobile trips, the size of the group traveling together (i.e., automobile occupancy). Door-to-door travel times between each city pair on each mode were estimated by aggregating 3 components: main-line time, from published schedules of the common carriers, and typical terminal access and delay times for each end of the trip. Fare and frequency of service data were tabulated by mode for each city pair.

To estimate the number of travelers likely to use high-speed rail service required the formulation of a set of high-speed rail service characteristics for each New York State corridor. These service characteristics included travel time, fare, frequency of service, and amenities of the high-speed rail service (e.g., food service, terminal parking, and attractiveness of equipment).

Door-to-door travel time via the high-speed train between each city pair was developed by adding typical ground access and terminal delay times for each end of the trip to the main-line (station-to-station) time. Station-to-station times were taken from a simulation of high-speed rail service between New York City and Buffalo developed by the United Aircraft Corporation (8). The simulation was made for the existing right-of-way by using a maximum speed of 120 mph, and it reflects all the physical restraints on train speed (e.g., tunnels, curves, and bridges). The resulting rail speeds represent an increase of approximately 50 percent over those of conventional rail service. Travel times between major cities are expected to be reduced as

follows: between New York City and Albany by approximately 1 hour, between New York City and Buffalo by approximately 3 hours, and between New York City and Montreal by approximately 4 hours.

High-speed rail fares were assumed to be in the general range of present rail fares—lower than air fares, but higher than bus fares. Pinpointing future fares exactly is not only very difficult but also unnecessary for this model; future fares depend on equipment costs, physical improvements, operating costs and revenues, competition, possible government support, and other economic factors.

The following frequency-of-service guidelines were adopted for this study. First, future high-speed rail service is expected to be more frequent than present rail service with its 2-hour minimum headway between New York City and Albany. The hourly service anticipated in the Northeast Corridor is viewed as a feasible upper limit of service frequency. Service is expected to be as frequent as that of alternative modes. Consequently, frequency of service has not been accounted for explicitly in estimating the potential passenger market for high-speed rail service. The amenities of high-speed rail service were assumed to be of high quality, comparable to those of air and high-speed rail service now available in the Northeast Corridor.

Estimation of Diversions From Air Travel

The model for estimating the number of air travelers likely to divert to high-speed rail service is shown in Figure 3. The branch diagram shown in Figure 3 illustrates how intercity air trips were stratified into market segments.

An analysis of air fares relative to high-speed rail fares indicated that air fares were greater than or equal to rail fares for all city pairs. This finding provided one of the 2 basic inputs needed to determine the diversion category of each market segment.

Because the model is applied to many city pairs, door-to-door travel time via high-speed rail service can be faster or slower than air service. For each travel-time condition, a market segment is classified into one of the 3 diversion categories by applying the reasoned decision rules mentioned previously. The reasons for classifying each market segment into a particular diversion category are also shown in the last column of Figure 3.

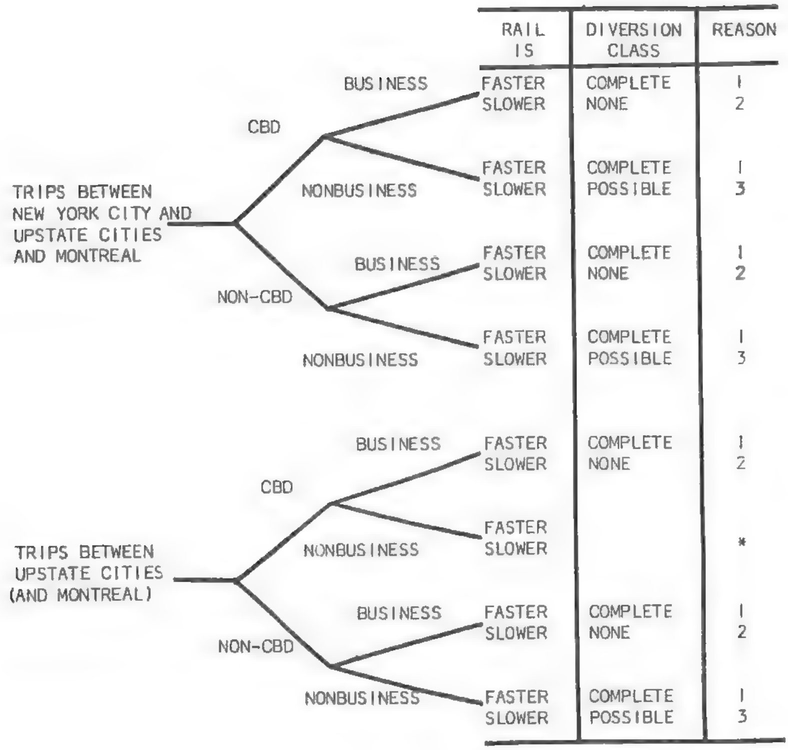
The set of diversion guidelines applied to air as well as to other trips originating or terminating in New York City was slightly different from the set applied to trips made solely between upstate New York cities. This refinement was made because of the considerably different terminal access characteristics in New York City relative to the smaller cities in upstate New York and because more detailed trip data were available for the New York City area.

In several instances, the diversion rules originally adopted were modified by judgment. For example, the non-CBD business market segment was considered nondivertible if the train was slower than air. Although this segment would have been considered as possibly divertible to high-speed rail service if the diversion classifications were strictly applied, it was classified as nondivertible because most business travelers would likely be willing to pay the premium fare for the time saved.

Estimation of Diversions From Automobile Travel

The model for estimating the number of automobile travelers likely to divert to high-speed rail service is given in Figure 4. Diversion guidelines for automobiles were developed in essentially the same manner as those for air travel, with 1 important difference. A comparison of the costs of intercity travel by automobile and by high-speed rail service showed that for a single-person trip the cost on each mode was essentially equal but that for a multi-person trip the cost of automobile travel was much less expensive than rail for all city pairs. This finding is reflected in the diversion guidelines.

In this study, 2-person automobile trips for business purposes to CBD and non-CBD locations in New York City were considered completely divertible if the train was faster than the automobile. Even though the train fare was more expensive than the



1. RAIL IS FASTER AND LESS EXPENSIVE THAN AIR, AND RAIL SERVICE IS SIMILAR IN QUALITY TO AIR SERVICE.

2. A BUSINESS TRAVELER NOW USING AIR IS NOT LIKELY TO DIVERT IF THE TRAIN IS SLOWER.

3. RAIL IS SLOWER BUT LESS EXPENSIVE THAN AIR, SO DIVERSIONS DEPEND ON HOW A NONBUSINESS TRAVELER VALUES HIS TIME.

* ALL AIR TRIPS TO CBD LOCATION IN UPSTATE NEW YORK CITIES AND IN MONTREAL WERE ASSUMED TO BE BUSINESS TRIPS.

Figure 3. Procedure used to estimate diversions from air to high-speed rail.

automobile in this case, it was assumed that the business traveler would be willing to pay the additional cost for reduced travel time.

Estimation of Diversions From Bus Travel

Figure 5 shows the model for estimating diversions from bus service to high-speed rail service. The guidelines for making these estimates were developed and applied in the same manner as those discussed for air and automobile trips, with 1 exception. All bus trips were assumed to originate and terminate in the CBD of the city pairs under study because both bus service and rail service are generally provided between these points. Differences in travel time and cost between bus and rail modes are considered to be a result of the line-haul portion of the trip and not of the ground-access portions.

There are no completely divertible bus passengers. Although high-speed rail service is generally faster than bus service, the latter is always less expensive. The number of diversions from bus to high-speed rail service will depend on how travelers

(a)

Trips between New York City and Upstate Cities and Montreal			Rail is	Diversion class	Reason
CBD	Business	1 person	FASTER	COMPLETE	1
			SLOWER	POSSIBLE	2
		2 persons	FASTER	COMPLETE	3
			SLOWER	NONE	4
		3 persons	FASTER	POSSIBLE	5
			SLOWER	NONE	4
	Nonbusiness	1 person	FASTER	POSSIBLE	6
			SLOWER	POSSIBLE	7
		2 persons	FASTER	POSSIBLE	6
			SLOWER	NONE	8
		3 persons	FASTER	NONE	9
			SLOWER	NONE	8
Non-CBD	Business	1 person	FASTER	COMPLETE	1
			SLOWER	POSSIBLE	2
		2 persons	FASTER	COMPLETE	3
			SLOWER	NONE	4
		3 persons	FASTER	POSSIBLE	5
			SLOWER	NONE	4
	Nonbusiness	1 person	FASTER	POSSIBLE	6
			SLOWER	NONE	10
		2 persons	FASTER	NONE	9
			SLOWER	NONE	8
		3 persons	FASTER	NONE	9
			SLOWER	NONE	8

1 RAIL IS FASTER AND ONLY SLIGHTLY MORE EXPENSIVE. A BUSINESSMAN WILL AVOID CBD AUTO CONGESTION AND CAN WORK ON THE TRAIN.

2 RAIL IS SLOWER AND SLIGHTLY MORE EXPENSIVE. HOWEVER, CBD AUTO CONGESTION IS AVOIDED AND WORK CAN BE DONE ON THE TRAIN.

3 RAIL IS FASTER AND MORE EXPENSIVE BUT OFFERS THE BUSINESSMAN THE SAME ADVANTAGES NOTED IN REASON 1.

4 MULTI-PERSON BUSINESS TRIPS ARE UNLIKELY TO DIVERT IF RAIL IS SLOWER AND MORE EXPENSIVE.

5 RAIL IS FASTER, BUT SIGNIFICANTLY MORE EXPENSIVE. HOWEVER, RAIL OFFERS THE BUSINESSMAN THE SAME ADVANTAGES NOTED IN REASON 1.

6 RAIL IS FASTER AND MORE EXPENSIVE. THE NUMBER OF DIVERSIONS DEPENDS ON HOW NONBUSINESS TRAVELERS VALUE THEIR TIME.

7 RAIL IS SLOWER AND SLIGHTLY MORE EXPENSIVE. THE NUMBER OF DIVERSIONS DEPENDS ON HOW NONBUSINESS TRAVELERS VALUE THEIR TIME. AUTO CONGESTION IS AVOIDED.

(b)

Trips between Upstate Cities including Montreal			Rail is	Diversion class	Reason
CBD	Business	1 person	FASTER	COMPLETE	11
			SLOWER	POSSIBLE	12
		2 persons	FASTER	COMPLETE	13
			SLOWER	NONE	4
		3 persons	FASTER	NONE	9
			SLOWER	NONE	4
	Nonbusiness			*	
Non-CBD	Business	1 person	FASTER	POSSIBLE	14
			SLOWER	NONE	10
		2 persons	FASTER	POSSIBLE	14
			SLOWER	NONE	10
		3 persons	FASTER	NONE	9
			SLOWER	NONE	8
	Nonbusiness	1 person	FASTER	POSSIBLE	14
			SLOWER	NONE	10
		2 persons	FASTER	NONE	9
			SLOWER	NONE	8
		3 persons	FASTER	NONE	9
			SLOWER	NONE	8

* THE NUMBER OF INTERCITY AUTO TRIPS MADE FOR NONBUSINESS PURPOSES TO THE CBD'S OF UPSTATE CITIES AND MONTREAL WERE ASSUMED TO BE NEGLIGIBLE.

8 RAIL IS SLOWER AND SIGNIFICANTLY MORE EXPENSIVE FOR A MULTI-PERSON TRIP.

9 RAIL IS FASTER, BUT SIGNIFICANTLY MORE EXPENSIVE FOR A MULTI-PERSON TRIP.

10 RAIL IS SLOWER AND MORE EXPENSIVE. CONGESTION IS LESS OF A PROBLEM.

11 RAIL IS FASTER AND ONLY SLIGHTLY MORE EXPENSIVE. A BUSINESSMAN CAN WORK ON THE TRAIN.

12 RAIL IS SLOWER AND SLIGHTLY MORE EXPENSIVE. HOWEVER, A BUSINESSMAN CAN WORK ON THE TRAIN.

13 RAIL IS FASTER AND MORE EXPENSIVE, BUT OFFERS THE SAME ADVANTAGES DESCRIBED IN REASON 1.

14 RAIL IS FASTER BUT SLIGHTLY MORE EXPENSIVE. AUTO TRAVELERS DESTINED TO NON-CBD LOCATIONS MAY BE WILLING TO DIVERT TO RAIL.

Figure 4. Procedure used to estimate diversions from automobile to high-speed rail (a) with New York City as 1 end of the trip and (b) for all other city pairs excluding New York City.

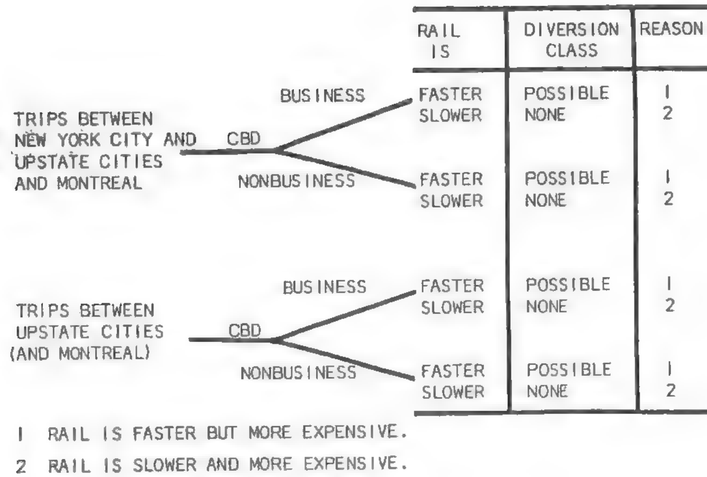


Figure 5. Procedure used to estimate diversions from bus to high-speed rail.

evaluate the trade-off between time and cost. Consequently, the estimated number of diversions from bus to rail is likely to be conservative, particularly because certain market segments, such as business travelers, may be attracted to improved rail service.

Estimation of Diversions From Existing Rail Service

All travelers using existing rail service in the 3 New York State corridors are expected to divert to high-speed rail service because this service is expected to be faster and more frequent than conventional rail service and similar in cost. It was assumed, when estimating high-speed rail ridership in a particular corridor (e.g., New York City-Albany), that conventional rail service would continue in adjoining corridor segments (e.g., Albany-Buffalo, Albany-Montreal). Consequently, transfers from conventional rail to high-speed rail service (but no transfers from other modes) were expected to occur between corridor segments at Albany—the city common to all corridors. The number of such transfers is noted in sample ridership estimates given in the next section.

HIGH-SPEED RAIL USE FOR 1968—AN APPLICATION OF THE MODEL

The diversion model was used to estimate high-speed rail use for 1968 for the 3 proposed rail corridors in New York State (Table 2). The high-speed rail service characteristics used in this test are those described in the previous discussion of the diversion model.

In terms of completely divertible travelers, high-speed rail ridership is highest in the New York City-Buffalo corridor and lowest in the New York City-Albany corridors. The same ranking of corridors is indicated by the possibly divertible ridership figures.

Approximately 60 percent of the completely divertible riders in each corridor are automobile travelers, approximately 35 percent are travelers using conventional rail service, and approximately 6 percent are air travelers. All bus passengers were classified as possibly divertible because, although high-speed rail service would be faster than bus service, it was assumed that high-speed rail service would also be more expensive.

The complete diversion figures show that high-speed rail service would increase annual rail ridership by 1,010,000 trips (183 percent) in the New York City-Buffalo

TABLE 2

ESTIMATED 1968 HIGH-SPEED RAIL RIDERSHIP AND TRAVEL VOLUMES

Item	New York City-Buffalo	New York City-Montreal	New York City-Albany
High-speed rail ridership by corridor[a]			
Completely divertible travelers from			
Existing rail, within corridor	471,000	232,000	134,000
Existing rail, transfer	80,000	209,000	289,000
Air	106,000	55,000	54,000
Automobile	904,000	542,000	497,000
Bus	0	0	0
Total	1,561,000	1,038,000	974,000
Possibly divertible travelers from			
Existing rail	0	0	0
Air	698,000	337,000	51,000
Automobile	2,385,000	1,072,000	985,000
Bus	1,220,000	874,000	534,000
Total	4,303,000	2,283,000	1,570,000
Total travel			
Existing rail	551,000	441,000	423,000
Air	2,511,000	2,366,000	2,355,000
Automobile	16,032,000	11,605,000	8,895,000
Bus	1,368,000	1,061,000	958,000
Total	20,462,000	15,473,000	12,631,000

[a]High-speed rail ridership in the New York City-Albany corridor is included in the ridership estimates for both other corridors.

corridor, by 597,000 trips (135 percent) in the New York City-Montreal corridor, and by 551,000 trips (130 percent) in the New York City-Albany corridor. For these same corridors, rail ridership would increase by an additional 4,303,000, 2,283,000, and 1,570,000 trips if all trips in the possibly divertible category were to use high-speed rail service.

Approximately 7 percent of all intercity trips in each corridor are classified as completely divertible to high-speed rail service (Table 3). If all the possibly divertible travelers were also to use the train, they would constitute an additional 21.0 percent of all intercity trips made in the New York City-Buffalo corridor, 14.8 percent in the New York City-Montreal corridor, and 12.4 percent in the New York City-Albany corridor.

As the sample results show, the diversion model provides a range of estimated high-speed rail ridership when applied to a potential corridor. The minimum level of estimated ridership in each corridor is the number of completely divertible travelers; the maximum level is the sum of complete and possible diversions. However, only a small portion of the possibly divertible travelers are expected to divert to high-speed rail service. The model does not provide an estimate of the exact size of this portion.

DISCUSSION OF THE DIVERSION MODEL

A "reasoned" diversion model was used to estimate high-speed rail use because there is some question as to whether a statistically calibrated modal-split model

TABLE 3

ESTIMATED 1968 HIGH-SPEED RAIL RIDERSHIP AS A PERCENTAGE OF TOTAL INTERCITY TRAVEL

Diversion Class	New York City-Buffalo (percent)	New York City-Montreal (percent)	New York City-Albany (percent)
Completely divertible	7.6	6.7	7.7
Possibly divertible	21.0	14.8	12.4

developed with travel volumes and service characteristics for existing rail service can accurately estimate high-speed rail use. This concern arises mainly because travel times, service frequency, and amenities of the proposed rail service are significantly different from those of existing rail service. It is possible that a traveler's implied weighting of travel time, cost, or service frequency in a mathematical model, based on present rail data, may not accurately represent the weighting of these factors for the proposed service.

A problem related to using a mathematical model based on data for existing rail service is that the model will likely have to be extrapolated beyond the range of data points used for calibration. For example, if the ratio of rail to automobile travel time is used as an independent variable in a model, the numerical values of this variable would likely change from 1.0 or slightly greater for present rail service to as low as 0.5 for the proposed rail service. Service frequency ratios and other travel time ratios, which are commonly used independent variables in modal-split models, would also change significantly for high-speed rail service. A potential solution to these problems is the use of an attitudinal modeling approach to estimate high-speed rail use.

Although the diversion model is elementary, it permits estimates to be made of high-speed rail use as a function of 5 important factors affecting modal split. Furthermore, the model produces reasonable preliminary estimates of high-speed rail use. Another reason for using this model in preliminary studies of high-speed rail service is that it does not have to be statistically calibrated. This reduces the time and expense of conducting a preliminary feasibility study.

An important property of the model is that it distinguishes between those travelers likely to be attracted to high-speed rail service as opposed to those travelers who may be attracted. This factor, together with its simplicity, makes the model more readily understood by decision-makers and the general public, an important consideration when the feasibility of new or improved forms of transportation is studied.

In the New York State study, the diversion model was not used to estimate future high-speed rail use. However, it could be used to prepare such estimates, provided that the necessary inputs to the model are projected into the future.

The formulation of the diversion model also has certain weaknesses. It does not account for travel generated or induced by the provision of high-speed rail service. Although these trips are likely to be small in number relative to diverted trips, they may be important in establishing the feasibility of high-speed rail service.

Another limitation of the model is that a range rather than a specific estimate of potential ridership is determined because the model has no explicit mechanism to trade off travel time and travel cost for the possibly divertible travelers. The problem can be minimized if a reasonable set of decision rules are formulated for various travel-time and travel-cost conditions and applied to the possibly divertible travelers. Decision rules of this type were not used in the New York State study in order to minimize subjectivity in the high-speed rail ridership estimates.

Frequency of service and amenities of the various modes are not explicitly accounted for in this diversion model. The model accounts only indirectly for these factors by assuming that they are equal for all common-carrier modes.

CONCLUSIONS

The diversion model presented here can be used to prepare preliminary estimates of high-speed rail use in intercity travel corridors. The approach outlined in this paper is particularly applicable both to initial feasibility studies of proposed high-speed rail service and to studies that attempt to identify corridors in which high-speed rail service may be justified.

High-speed rail use is estimated as a function of travel time, travel cost, trip purpose, destination location, and, for the automobile trips, group size of intercity trips. The diversion model is a set of reasoned decision rules that categorize travelers using modes of transportation other than high-speed rail as completely divertible, possibly divertible, or nondivertible to high-speed rail service. The model has been used successfully in a study of proposed high-speed rail service in New York State.

REFERENCES

1. Rossi, L. P., and DiRenzo, J. F. High-Speed Rail Service in New York State: Concept Study. New York State Department of Transportation, Albany, Publ. FR070601, 1970.
2. Rossi, L. P., and DiRenzo, J. F. Preliminary Market Study of the Potential for High-Speed Rail Service in New York State. New York State Department of Transportation, Albany, unpublished, 1969.
3. Feasibility of High-Speed Rail Service. System Analysis Research Corp. and Thomas K. Dyer, Inc., Cambridge, Mass., 1969.
4. Fertal, M. J., et al. Modal Split—Documentation of Nine Methods for Estimating Transit Usage. Bureau of Public Roads, U.S. Department of Commerce, Dec. 1966.
5. Studies in Travel Demand, 3 Vols. Mathematica, Princeton, New Jersey, 1965, 1966, 1967.
6. Intercity Travel in the Northeast Corridor—1966. Office of High-Speed Ground Transportation, U.S. Dept. of Transportation, mimeographed, 1969.
7. Third Report on the High-Speed Ground Transportation Act of 1965. Secretary of Transportation. U.S. Department of Transportation, 1969.
8. Turbotrains for the New York Central System. Sikorsky Aircraft, Stratford, Conn., 1967.

ANALYSIS OF AIR PASSENGER TRAVEL IN THE TWIN CITIES METROPOLITAN AREA

Francis P. D. Navin and Richard P. Wolsfeld, Jr.,
Bather-Ringrose-Wolsfeld, Inc., Roseville, Minnesota

This paper documents the socioeconomic and travel characteristics of the air passenger and quantifies the relationships of several variables to the air passenger trip generation within the Twin Cities area. The results of the analysis may then be used as input in determining development policy for future locations of new major airport facilities.

•THE NUMBER of air passenger trips in the United States has increased 3 times in the past 10 years to over 168 million trips in 1969. Forecasts are that air trips will continue to increase at the same rate until 1980. Besides this increase in the absolute number of air passenger trips, urban trips related to other airport uses are also increasing. Air passenger and air passenger-related trips will be concentrated in time and space because of larger aircraft and an increased demand for air travel.

Aircraft landing requirements have necessitated an increase in the space requirements of major regional airports to approximately 5,000 acres. Control over noise, air pollution, and land development are important locational factors for regional airports. As a result, new regional airports will have to be located farther from the urban areas in order to satisfy stringent locational requirements. To plan adequate facilities for the air traveler within the airport complex requires that more be known about air traveler characteristics. The Twin Cities area was used as a model in an analysis of the problem (1).

AIR PASSENGERS

Air passengers make approximately 30 to 40 percent of all daily trips to and from air terminals (2, 3). An in-flight survey of commercial aircraft entering and leaving the Twin Cities, conducted during the week of January 21-28, 1970, showed that a total of 91,996 air passengers traveled to or from the Twin Cities. These 91,996 air passenger trips were divided into various air terminal movements: 32,691 terminated in the Twin Cities Region, 20,049 transferred to other planes, 5,111 passed through on the same aircraft, and 34,145 originated in the Twin Cities region.

The 66,836 regional trips in the Twin Cities area were further divided into those trips inside of the study area, defined by 756 traffic analysis zones for which socioeconomic data were available, and those trips outside of the study area. During the survey week, 57,839 or 86.5 percent of the regional trips originated or terminated within the study area.

COMBINING ARRIVING AND DEPARTING PASSENGERS

The initial step in the analysis was to determine whether arriving and departing passengers had similar socioeconomic characteristics (4). If the characteristics were similar, then the arriving and departing passenger groups could be combined for further analysis. The respondents to the air-travel interview were categorized into 4 principal groups that were as follows: arriving residents, departing residents, arriving visitors, and departing visitors.

Sponsored by Committee on Transportation Forecasting and presented at the 50th Annual Meeting.

These groups were further classified by the following characteristics: purpose of air trip, address type at origin or destination, mode of ground travel, and family income. These categories and cross classifications were then used to establish the similarity of arriving and departing passengers. The following discussions of the address type and income relationships is an example of the similarity that exists within the total air passenger population.

Address Type

The differences in address type between residents and visitors who arrive and depart from the airport are given in Table 1. Residents had the most trip ends at homes, whereas visitors had trip ends usually at hotels or motels. Visitors were also more likely to have trip ends at places of business than were residents. The comparison between arriving and departing passengers also indicated differences in distribution of address types. For residents, departing passengers were more likely to have their origin trip ends at their places of employment than those arriving. More than 90 percent of all arriving residents had trip ends at their homes as compared to 84 percent for departing residents.

For visitors, the greatest difference in land use categories for arriving and departing passengers was between hotels or motels and places of business. Arriving passengers were more likely to have their trip ends at hotels or motels than were departing passengers. However, departing visitors usually left from places of business. None of the other differences between arriving and departing visitors exceeded 2 percent. Therefore, as in the case of residents, pooling the distributions of land use categories for arriving and departing visitors did not distort the importance of the different land use categories.

Income

As given in Table 1, the maximum difference for any income category occurred in the $15,000 to $19,999 group and ranged from a low of 20.0 percent for arriving residents to 23.2 percent for departing visitors. All other differences among the 4 types of passengers were less than 3 percent. Even though residents and visitors had different characteristics when categorized by trip purpose, address type at origin or destination, and mode of travel, there was no evident reason to expect income distribution to be different for residents and visitors.

Number of Models Required

A summary of the resident and visitor air trips by the 2 major divisions of address type is given in Table 2. The residents have 90 percent of their trip origins at homes. It was assumed that the characteristics of the zone could represent the characteristic of the air passenger. An investigation indicated a similarity between the resident and visitor air passenger whose trips were home based; therefore, these 2 groups were combined.

More than 50 percent of the visitors' trips were associated with hotels or motels. The majority of remaining visitor trips were associated with non-home addresses. Because visitors with home-based trips were shown to have the essential characteristics of residents, the model for non-home-based trips could be based on the characteristics of hotels or motels rather than other land use characteristics.

Summary

Studying the various socioeconomic characteristics of all the air passengers showed that arriving and departing passengers were similar. The similarities allowed a pooling of the trips into larger groups. An analysis of these groups showed that essentially 2 models would explain the air passenger trip-generation characteristics: one for home-based air passenger trips, which accounted for 90 percent of all resident trips, and one for non-home-based air passenger trips, which accounted for 80 percent of all the visitor trips.

TABLE 1

AIR PASSENGER TRIPS BY INCOME AND ADDRESS TYPE

Income	Arriving							Departing						
	Home or Other Than Home	Hotel or Motel	Normal Work Place	Other Than Work Place	Other	Total	Percent of Total	Home or Other Than Home	Hotel or Motel	Normal Work Place	Other Than Work Place	Other	Total	Percent of Total
Residents														
$0-4,999	397		4		32	433	3.2	403	18	24		23	468	3.0
$5,000-9,999	1,134	14	77		40	1,265	9.1	1,481	10	40		88	1,619	10.3
$10,000-14,999	2,587	45	166		36	2,834	20.5	2,739	30	376	20	104	3,269	20.8
$15,000-19,999	2,490	31	196	4	38	2,759	20.0	2,821	47	451	18	66	3,403	21.6
$20,000-29,999	2,452	30	107	8	15	2,612	19.9	2,459	12	509	14	17	3,011	19.2
More than $30,000	2,269	39	176	12	34	2,530	18.3	2,100	53	360	31	51	2,595	16.5
Not given	1,280	20	56		28	1,384	10.0	1,160	19	136	3	40	1,358	8.6
Total	12,609	179	782	24	223	13,817		13,163	189	1,896	86	389	15,723	
Percent of Total	91.2	1.3	5.7	0.2	1.6		100.0	83.7	1.2	12.1	0.5	3.5		100.0
Visitors														
$0-4,999	221	68		15	99	403	3.1	271	123	10	10	71	485	3.8
$5,000-9,999	459	472	17	91	216	1,255	9.7	396	420	23	73	223	1,135	8.8
$10,000-14,999	468	1,398	38	359	284	2,547	19.8	592	1,166	50	432	263	2,503	19.5
$15,000-19,999	311	1,564	78	460	247	2,660	20.7	475	1,514	122	578	7,288	9,977	23.2
$20,000-29,999	235	1,619	17	448	211	2,530	19.6	265	1,332	84	583	173	2,437	19.0
More than $30,000	223	1,491	33	400	303	2,378	18.4	304	1,278	62	475	259	2,378	18.5
Not given	237	501	32	164	184	1,118	8.7	116	414	36	144	167	927	7.2
Total	2,154	7,108	215	1,937	1,477	12,891		2,469	6,247	387	2,295	1,444	12,842	
Percent of Total	16.7	55.1	1.7	15.0	1.5		100.0	19.3	48.5	3.0	17.9	11.3		100.0

TABLE 2

AIR PASSENGER TRIPS

Passenger	Trips		Total
	Home-Based	Non-Home-Based	
Residents	25,772	3,767	29,539
Visitors (hotel or motel)	0	13,354	13,354
Visitors (other)	4,624	7,757	12,381
Total	30,396	24,878	55,274

NONSPATIAL CHARACTERISTICS OF AIR PASSENGERS

The aggregate characteristics for arriving and departing travelers were grouped as follows: air trip purpose, address type at origin or destination, family income, and modes of ground travel.

Air Trip Purpose

Purposes of air trips for all passenger trips in the Twin Cities area are shown in Figure 1. Most trips were for business purposes, 33,827 trips, which were distributed equally between residents and visitors. "Manufacturing and wholesaling" and "personal or professional service" accounted for 63 percent of all business trips. The trip-purpose category with the next highest number of trips, 7,919, was "personal." One other interesting fact was that the January vacation trips of residents outnumbered the visitors by more than 3 to 1.

Address Type at Origin or Destination

Eighty-five percent of residents, or 45 percent of all air trips, started or ended trips at their homes. The origins or destinations of the 17,645 trips by visitors were from or to hotels or motels or someone else's home. Visitors tended to have more trip ends at hotels and more origins at other work places.

Family Income

The majority of people who flew during the test week had an annual family income more than $10,000. Individuals with incomes of more than $20,000 accounted for 40 percent of the air patronage.

The number of air trips made per year also increased with income. Those passengers who made 1 trip per year were twice as likely to have incomes of more than $10,000 than to have incomes of less than $10,000. Passengers were 7 times more likely to be in the over-$10,000 income group than to be in the under-$10,000 income group if they made 8 to 12 flights per year and 10 times more likely to be in the higher income group if they made more than 21 flights per year. When the income characteristics were cross correlated with trip purposes, the majority of low-income passengers were found to be either military personnel or students.

Modes of Ground Travel

Air passengers who drove their own cars or who were driven by someone else numbered 34,511 or constituted 62 percent of all ground trips. The second largest group of passengers who used taxis, airport buses, and limousines, accounted for an additional 25 percent of the trips. Car rental transportation was used by 4,764 air passengers, most of whom were visitors. Courtesy cars and public buses were used for 1.2 percent of all the ground trips made by airline patrons.

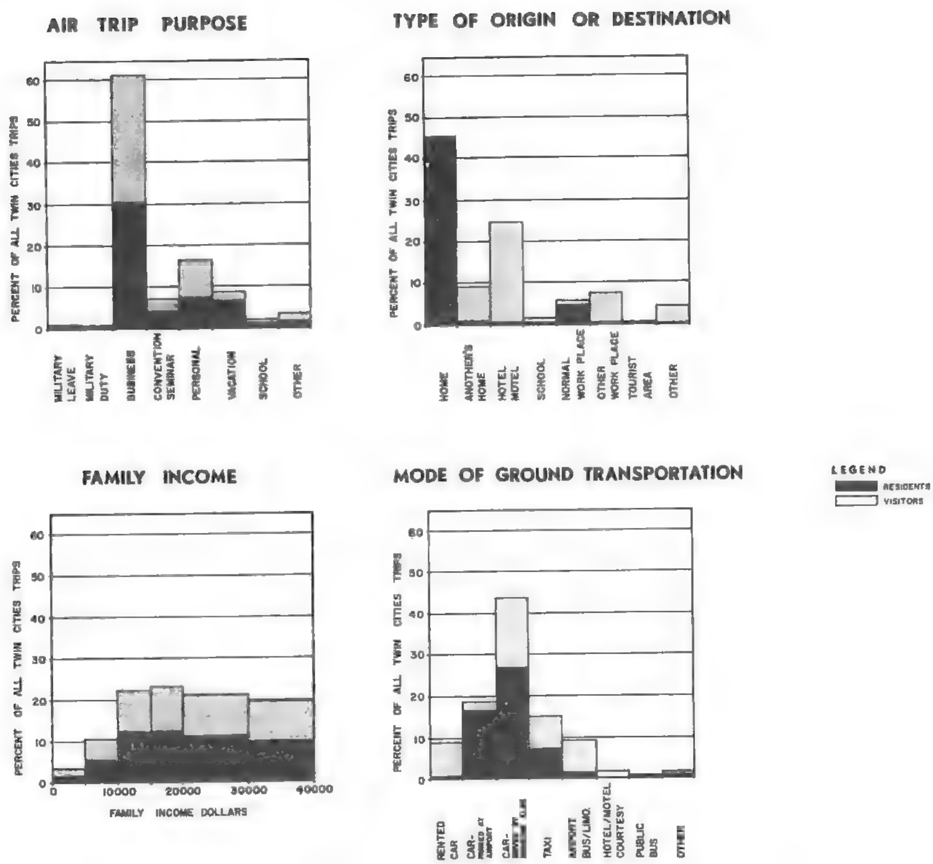


Figure 1. Travel characteristics of air passengers at Twin Cities Airport.

TRAVEL TIME CHARACTERISTICS OF AIR PASSENGERS

The average travel time for all air passengers, as shown in Figure 2, was 25.5 min. The travel times from each traffic analysis zone in the study area to the air terminal were obtained from the Minnesota Department of Highways computerized 1970 highway network system. The trip-length distribution for various types of trips is shown in Figure 2 and discussed in the following sections.

All Home-Based Trips

There were 30,396 air passenger trips, or 55 percent of all trips, that had a home base at 1 end. Within 10 min driving time from the air terminal, there were 1,400 air passenger trips, and this number increased at an average rate of 1,000 trips per minute for the next 20 min. The number of trips then decreased until at 50 minutes only 4 percent of air trips remained. Twenty-five percent of all home-based trips were within a 20-min travel time and 75 percent were within 34-min travel time to the terminal. The average trip for all home-based trips was 27.7 min.

Total Resident Trips

The distribution of travel time for total resident trips was generally similar to that for all home-based trips. It rose to a maximum of 1,500 trips at 22 min and then decreased to almost 0 trips at 50 min. At 22 min, 42 percent of the residents had

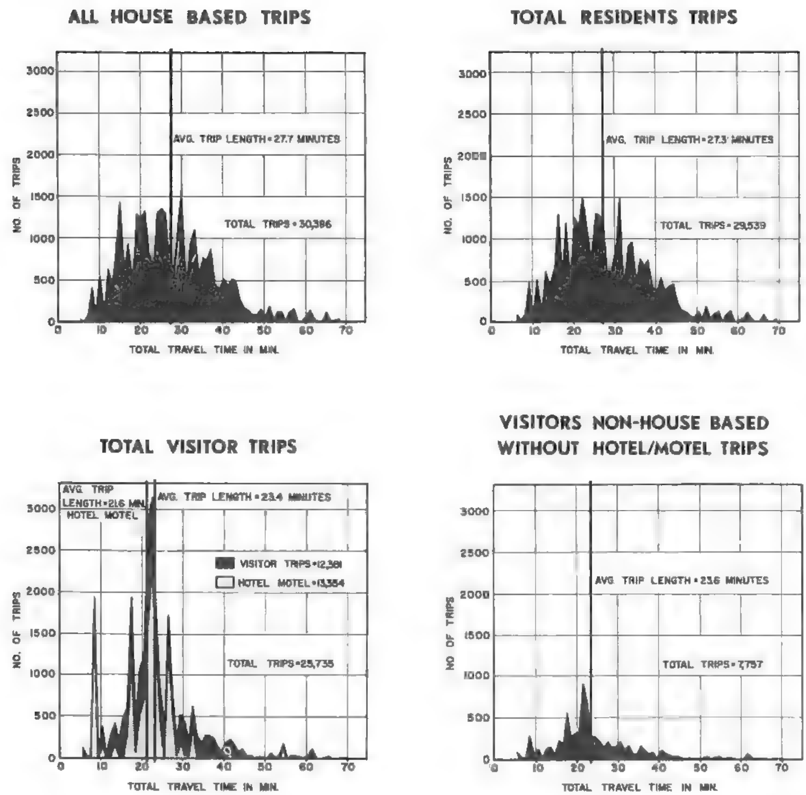


Figure 2. Travel time characteristics of air passengers at Twin Cities Airport.

completed their ground trips, and, at 50 min, 4 percent had yet to complete their trips. The average travel time for this set of air patrons was 27.3 min.

All Non-Home-Based Trips

The 7,757 non-home-based visitor trips increased to the peak of approximately 900 trips at 22 min. Beyond the peak, there was a rapid drop to a few hundred trips, and then a very gradual decrease to almost 0 trips at the outer reaches of the study area. The average travel time was 23.6 min, with 25 percent of all the trips being within 17 min and 80 percent being within 31 min of the terminal.

Total Visitor Trips

The major hotel areas adjacent to the airport were 8 min and the 2 city centers of St. Paul and Minneapolis were at 23 and 18 min respectively from the terminal. The remaining visitors trips were regularly distributed. They increased to a maximum of 23 min from the air terminal and then slowly decreased. The average trip time for the 13,354 hotel-based visitor trips was 21.6 min, and 90 percent of all hotel trips were within 27 min travel time. Of all remaining visitor trips, 25 percent were within 30 min, and the average trip time was 23.4 minutes.

AIR PASSENGER TRAVEL MODEL

The second objective of the analysis was the development of mathematical models to simulate the previously described travel patterns. The trip generation model

estimates the number of air passenger trip ends in each traffic analysis zone based on the existing socioeconomic characteristics and travel to the airport. The equations are then used to estimate future air passenger origins and destinations throughout the metropolitan area on the basis of the forecast characteristics of the zone and a changing propensity of the population to use the air mode.

The variables used in developing the model must be those for which forecasts are available and reliable. The zonal variables with reliable 1985 forecasts in the Twin Cities are median family income, number of people, number of dwelling units, number of employees by selected industrial categories, and number of hotel or motel rooms. The characteristics of the air passengers are purpose of trip, address type, travel time to the airport, and income as previously indicated.

The minimum number of classes of air passengers, as previously explained, were determined to be a set of two—all home-based trips and all non-home-based trips. The resulting equations were obtained by a stepwise linear regression that used socioeconomic variables that were logically and statistically significant. The equations were then corrected for spatial distribution with the addition of a travel time function.

First Selection of the Important Independent Variables

As a first step in selecting which of the 13 independent variables to use in the equations, a simple correlation matrix was developed. The coefficient r is a measure of association between 2 variables, and the matrix gives the correlation coefficients for all combinations of variables. The existence of a strong relationship indicates a possible usable variable for the model. Investigation of the simple correlation matrix and numerous stepwise regression runs of home-based, non-home-based, and hotel- or motel-based air trips against all other variables indicated the following:

1. Non-home-based trips are highly related to the number of hotel or motel rooms (the best indicator) and to employment characteristics. A linear relationship was noted between these characteristics and non-home-based trips.
2. Two zones generated non-home-based trips that were not significantly related to the urban characteristics of the zone. These were motel areas adjacent to the southwest corner of the airport and the St. Paul CBD.
3. Areas farther from the airport have lower generation rates than those that are closer to the airport.
4. Home-based trips relate to economic groups. The zones where the median income was less than \$11,000 had low air trip-making rates and zones where it was above \$21,000 had high air trip-making rates.

Non-Home-Based Model

Stepwise regression analysis yielded the statistics for the non-home-based model given in Table 3. Dropping those variables with a t-value of less than 2.0 yielded the following equation:

$$T_{NHB} = 0.0108\ WM + 0.104\ R + 0.0311\ F + 0.0346\ GS + 1.145\ HM + 5.123$$

Home-Based Model

Running the standard stepwise regression to yield a model for home-based trips provided low coefficients of correlation. When the calculated values were compared with observed values, the low-income areas were estimated to be high and the high-income areas were estimated to be low. The next step was to compare the trip-making rates (air trips per 100 dwelling units) versus the median income for each traffic analysis zone. Then, the application of a modeling technique (5) that related trip length, density of trips, and socioeconomic variables would give the generalized relationship shown in Figure 3.

At least squares fit to each of the income ranges (\$0-11,000, \$11,000-21,000, and more than \$21,000) resulted in the following equations:

$$T_{HB} = 2.5\ (DU/100)\ \text{for}\ I < \$11{,}000$$

$$T_{HB} = (-6.8 + 0.00089)\ I\ (DU/100)\ \text{for}\ \$11{,}000 \leq I \leq \$21{,}000$$

$$T_{HB} = 11.0\ (DU/100)\ \text{for}\ I > \$21{,}000$$

where

T_{HB} = home-based air passenger trips;
I = median family income in the zone, dollars; and
DU = number of dwelling units.

TABLE 3

NON-HOME-BASED MODEL VARIABLES

Variable	t-Value	r^2	Δr^2	Coefficient
Hotel or motel, HM	35.2	0.9800		1.145
Government service, GS	16.4	0.9857	0.0057	0.0346
Wholesale and manufacturing, WM	6.2	0.9869	0.0012	0.0108
Finance, insurance, and real estate, F	2.3	0.9879	0.0010	0.0311
Retail, R	2.2	0.9872	0.0007	0.104
Education and local government	1.2			
Transportation, utilities, and community	1.4			
Personal and professional service	1.1			

Travel Time Function

The ratio of the number of observed air trips to the number of calculated air trips was plotted against travel time to the airport for each traffic analysis zone for both home-based and non-home-based trips. A least squares fit was made to these points as shown in Figure 4. These travel time functions were incorporated into the model as a factor to be multiplied by the original equation.

K-Factors

As previously mentioned, zonal adjustments for 2 zones, the St. Paul CBD and a hotel or motel area near the airport, were necessary because of the orientation between the activities of the zones and those of the airport. The K-factor for the St. Paul CBD, which was 0.55, is chiefly explained by the nonairport use of the hotel or motel rooms. The motel zone close to the airport had a K-factor of 2.0.

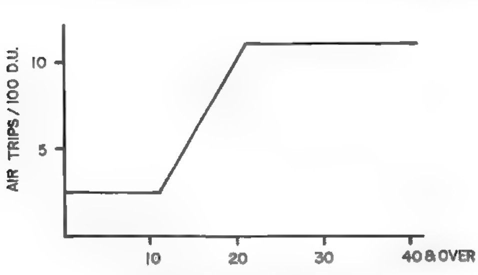


Figure 3. Trip generation model for air passenger trips originating from home.

Final Equations

The final trip-generation equations for existing conditions with the travel time function and K-factors were as follows:

1. Non-Home-Based Equations

$$T_{NHB} = (0.0108\ WM + 0.0104\ R + 0.0311\ F + 0.0346\ GS + 1.145\ HM + 5.123)\ (1.565 - 0.0217\ tt)\ K_j$$

where

K_j = 2.00 for zone 176, and
K_j = 0.55 for the St. Paul CBD.

2. Home-Based Equations

$$T_{HB} = 2.5\ (1.26 - 0.005\ tt)\ (DU/100)\ \text{for}\ I < \$11,000$$

$$T_{HB} = (-6.8 + 0.00089)\ I\ (1.26 - 0.005\ tt)\ (DU/100)\ \text{for}\ \$11,000 \leq I \leq \$21,000$$

$$T_{HB} = 11.0\ (1.26 - 0.005\ tt)\ (DU/100)\ \text{for}\ I > \$21,000$$

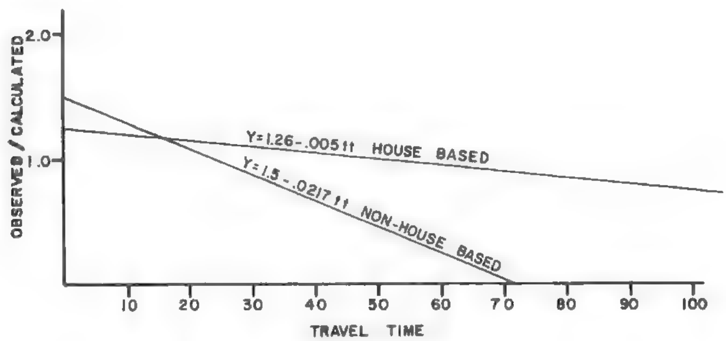


Figure 4. Air passenger travel time function.

Applying these equations to the base-year socioeconomic data resulted in the statistics given in Table 4. The low coefficient of correlation for the home-based trips resulted from the finding that the income of the air passenger was only weakly associated with the mean zonal income. The mean zonal income or some measure of the distribution of zonal income may yield a stronger relationship. The resulting non-home-based trip relationship was considered to be well within the limits of accuracy for trip generation.

APPLICATION OF THE MODEL

The air passenger accessibility study was one of many technical studies undertaken to help determine the most suitable location for a possible new air terminal for the Twin Cities (6). The 2 sites under consideration are shown in Figure 5. The northern site at Ham Lake is 18.5 miles north of the cities, and Farmington is 20.0 miles to the south. This section explains the modifications to the developed travel models necessary to account for an increased propensity to use aircraft and a changing travel time function.

Adjustment to the Travel Time Function

The non-home-based travel time function for the base year was (1.56 - 0.022 travel time). As a result, at 26 min, its effect is neutral and at 72 min, it reduces the number of trips to 0. Because both the Ham Lake and the Farmington sites will have longer travel times compared to Wold Chamberlain, the model that explained the existing situation must be adjusted to reflect this increase in travel time. The coefficient of travel time was adjusted by the ratio of the difference between the sum of all travel times to the new site and the sum of all travel times to Wold Chamberlain divided by the sum of all travel times to the new site. Mathematically, this is given as

$$\text{Adjustment} = \frac{\sum_{i=1}^{756\ \text{zones}} \text{tt (new site)} - \sum_{i=1}^{756\ \text{zones}} \text{tt (existing site)}}{\sum_{i=1}^{756\ \text{zones}} \text{tt (new site)}}$$

TABLE 4

APPLICATION OF TRIP-GENERATION EQUATIONS TO BASE-YEAR SOCIOECONOMIC DATA

Variables	Trips	
	Home-Based	Non-Home-Based
Observed trips	30,077	25,088
Calculated trips	29,937	24,600
r^2	0.47	0.96

This was done for the home-based and non-home-based equations. The resultant factors were as follows:

Location	Sum of Travel Times	Adjusted Factors
Ham Lake	37,591 min	0.402
Farmington	33,507 min	0.328

Adjustment Due to an Increased Propensity to Air Travel

The increase in air travel between 1970 and 1985 will be resulting from increases in employment, population, median income, and participation. This increased participation will be by people of the same 1970 and 1985 socioeconomic groups who fly more and also by those in lower socioeconomic groups. The non-home-based model as developed only takes into account the increases due to employment, hotel or motel rooms, and the like. Thus, with no adjustment, this model indicates that, if a zone increases employment from 1,000 to 2,000, the zone will produce air trips in 1985 similar to a zone that had 2,000 employees in 1970. The home-based model considers increased participation in the \$11,000 to \$21,000 range because the number of air trips increased with increased income in this range. All increased participation was adjusted by running the models adjusted with 1985 socioeconomic and travel time data. The 1985 forecast of total air passenger departures numbering 11,265,800 was made independently (7). This 1985 aggregate was used to develop a forecast for air travel in and out of the study area for

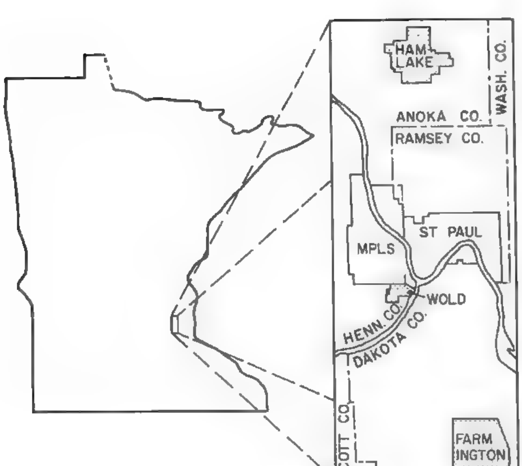


Figure 5. Existing and proposed major airport sites.

TABLE 5

1985 WEEKLY AIR TRIPS

Trips	Model-Unadjusted	Forecast	Ratio
Home-based			
Ham Lake	85,390	126,500	1.48
Farmington	87,415	126,500	1.45
Non-home-based			
Ham Lake	37,051	103,500	2.79
Farmington	44,774	103,500	2.31
Total	127,315	230,000	

the applicable week in January 1985. The ratio of the air travel found independently to the air travel obtained by the model was computed and applied to each coefficient. The corrections for increased participation are given in Table 5.

These 2 adjustments to the model made it possible to forecast the location and intensity of metropolitan area weekly air passengers for January 1985 for the Ham Lake and Farmington Airport sites.

External Air Travel

The preceding air passenger travel analysis was concerned with only the trips originating and terminating in the 756 zone study area. These trips accounted for 86.5 percent of all trips. The 8,997 air trips included in the area beyond the study area were treated in a different manner.

A growth factor technique was used to forecast the number and location of air trips in the outlying areas. If the assumption was true that the percentage of nonstudy area air trips in relation to all air trips remains constant over time, then the following relationships developed:

Air Trip End Location	Existing, 1970	Forecast, 1985
Study area	57,839	230,000
Outside of study area	8,997	35,780

Therefore, the nonstudy area air trips to or from Wold Chamberlain in the week of January 21, 1985, would be 35,780. Comparing this with the existing nonstudy area air trips yields a growth factor of 39.8 percent (35,780/8,997). This factor was multiplied by the existing air trips for each county on the assumption that distribution of air trips to the outlying counties would be similar to the existing distribution.

EVALUATION

The primary requirement of the results of the origin-destination survey and analysis was the comparison of the hours of travel by air passengers to each of the sites.

Study area travel

The air passenger hours of travel to each site are computed by multiplying the number of air passengers by the travel time to the particular airport for each traffic analysis zone and then summing the results for all zones.

The travel required for the forecast 230,000 study area weekly air passengers to reach the Ham Lake site was 184,063 hours. The travel required for the same number of air passengers to reach the Farmington site was 162,519 hours. Thus, 13.2 percent

more travel time would be required for the study area air passengers to reach the Ham Lake than to reach the Farmington site. The average trip length of the air traveler to Ham Lake was 48.0 min as compared to the average trip length of the air traveler to the Farmington site, which was 42.5 min.

Nonstudy area travel

The hours of travel required for the air passengers from the outlying area was computed in a similar manner. The travel times from each Minnesota and Wisconsin county that had air trips to the Ham Lake and the Farmington sites were computed. The forecast air trips for each county were then multiplied by these travel times to arrive at the hours of travel by the passengers to each site. The results indicated that 4,329 more hours of travel would be required for the Ham Lake site than for the Farmington site.

All Travel in Region

The combined additional hours of travel that would be expended by air passengers in 1985 to access the Ham Lake site versus the Farmington airport site was 25,873 hours. This represents a 14.2 percent difference in travel hours for the week of January 21, 1985, as compared to the same week in 1970.

CONCLUSIONS

The development of an air passenger trip-generation model requires first that the characteristics of the population using air travel be analyzed and then the number of models be reduced to the minimum. The next limiting factor is the availability of base-year and forecast socioeconomic characteristics. In addition, the variables should allow for a differentiated propensity to air travel by various economic and social groups. For example, the availability of income data (8) for each traffic analysis zone satisfies this requirement. Stratification of income and employment are the most important variables.

The total air passenger trips must be obtained through an independent analysis and will act as a control total to allow revision of the trip generation equations to future air travel trends. If the objective of the study is to evaluate competing sites for a new air terminal, then adjustments should be made to the travel time function to allow for a change in the travel function.

The trip generation techniques developed in more general transportation studies may successfully be applied to commercial air passenger travel.

ACKNOWLEDGMENT

The authors wish to thank Matthew Huber of the University of Minnesota and Thomas Ehlen formerly of Bather-Ringrose-Wolsfeld, Inc., for their assistance during the study. Assistance by Robert Einsweiler and other staff members at the Metropolitan Council is also gratefully acknowledged. The cooperation of the Office of System Planning of the Minnesota Department of Highways in providing the highway networks used in the analysis is appreciated.

REFERENCES

1. Twin Cities Airport Transportation Study. Bather-Ringrose-Wolsfeld, Inc., St. Paul, Minn., June 1970.
2. Keefer, L. E. Urban Travel Patterns for Airports, Shopping Centers, and Industrial Plants. NCHRP Rept. 24, 1966, 116 pp.
3. Air Transportation Analysis and Forecast. Regional Planning Commission, Franklin County, Columbus, Ohio, July 1969.
4. Corradino, J. C. The Philadelphia Airport Origin-Destination Survey—A Statistical Analysis. Highway Research Record 330, 1970, pp. 30-36.

5. Navin, F. P. D., and Schultz, G. W. A Technique to Calibrate Choice Models. Highway Research Record 322, 1970, pp. 68-83.
6. Newsletter. Metropolitan Council of the Twin Cities Area, Vol. 3, No. 5, St. Paul, Minn., July 1970.
7. Summary of Air Travel Demand Forecast for the Twin Cities Region. R. Dixon, Speas and Associates, working paper.
8. An Analysis of Airport Travel Demands, Lambert-St. Louis Municipal Airport. A. M. Voorhees and Associates, Inc., and Crawford, Bunte, Roden, Inc., Oct. 1969.

ESTIMATING MULTIMODE TRANSIT USE IN A CORRIDOR ANALYSIS

Gordon W. Schultz and Richard H. Pratt,
R. H. Pratt Associates, Garrett Park, Maryland

Experience is needed in the application of mode-choice modeling to detailed submode analysis, model calibration from limited surveys, and corridor-level planning. The paper describes an example of such applications. In addition, the example provides a partial test of previously postulated mode-choice theory. A chain of models based on the utilitarian theory of travel mode choice was used to reproduce and forecast use on competing bus and rail facilities in a suburban Chicago corridor. The models were calibrated by data from a survey area encompassing about a fifth of the study area population. Networks and trip tables were reduced from regional scope for corridor application. The corridor-level analysis allowed inexpensive testing of multiple-design alternatives. Model tests exhibited a satisfactory ability to reproduce data from study area transit rider counts and surveys, including mode shifts related to introduction of the Skokie Swift rapid transit service.

•PREDICTIVE MODELS for forecasting choice of travel mode are now in relatively common use for short-range as well as for long-range planning. Some of these applications involve study areas with inherently competitive transit operations and require analysis at the submode level of detail. At the same time, cost and study time limitations favor model calibration from small data sets if the resultant models can be shown to be reliable. Similarly, present trends favor planning techniques that can be applied to districts, corridors, and other subareas of the region.

These needs place a broad range of demands on the transportation analyst preparing to calibrate and apply mode-choice predictive models. In short, planners are now seeking choice models and procedures that can be applied at the submode level of detail, calibrated from limited data, and utilized effectively for subarea and corridor planning.

In a previous paper by Pratt (2), it was suggested that development of an underlying theory of travel mode choice would be a significant aid in meeting modeling requirements, including needs of this nature. A theory was set forth to the effect that individual choice of mode is utilitarian, that common measures of individual trip utility and choice are subject to chance errors describable by the normal distribution error function, and that deviations result from predictable influences. Some limited applications of the theory were presented, each of them utilizing the modeling technique of describing travel time, convenience, and cost with a single utility measure.

The work described in this paper moves a step forward by involving new application of the theory. Of particular interest is that this added application has involved study of competitive transit services, use of partial survey coverage, and analysis within a single corridor. Thus, the study described has provided experience both in general use of the theory previously described and more specifically in use of the theory for furthering detailed submode analysis, calibration from limited data, and corridor-level planning.

Sponsored by Committee on Transportation Forecasting and presented at the 50th Annual Meeting.

MODEL REQUIREMENTS OF THE STUDY

Area Description

The work that provides the basis for this case study was performed in the course of a public transportation study for the North Suburban (Chicago) Transportation Council (1). The project was financed by means of a UMTA Technical Studies Grant and by contributions from suburban villages and municipalities. Technical supervision was provided by the Chicago Area Transportation Study (CATS).

The study area covers a band along Lake Michigan roughly 7 miles wide. This corridor extends from the Chicago city limits north past the Cook County line to the Lake County communities of Libertyville and Lake Bluff. Figure 1 shows the study area, its rail services, and the home-interview survey area.

In the corridor, bus transit service is provided by 3 private companies and the Chicago Transit Authority (CTA). Commuter railroad and rail rapid transit service to Chicago is furnished by the Chicago and Northwestern Railway, the Milwaukee Road, and CTA.

In the southern portion of the 7-mile wide corridor, there are 4 roughly parallel rail services to Chicago. These are the Evanston Rapid Transit Line of CTA, the North Line of the Chicago and Northwestern Railway, the Skokie Swift Rapid Transit Shuttle of CTA, and the North Line of the Milwaukee Road. In addition, the Kennedy and Ravenswood CTA rapid transit services terminate close to the southern boundary of the study area.

These various transit alternatives offered to study area commuters differ not only in speed but also in cost, riding comfort, and frequency of service. Needless to say, there is extensive interaction between the private right-of-way transit lines as they affect rider choice.

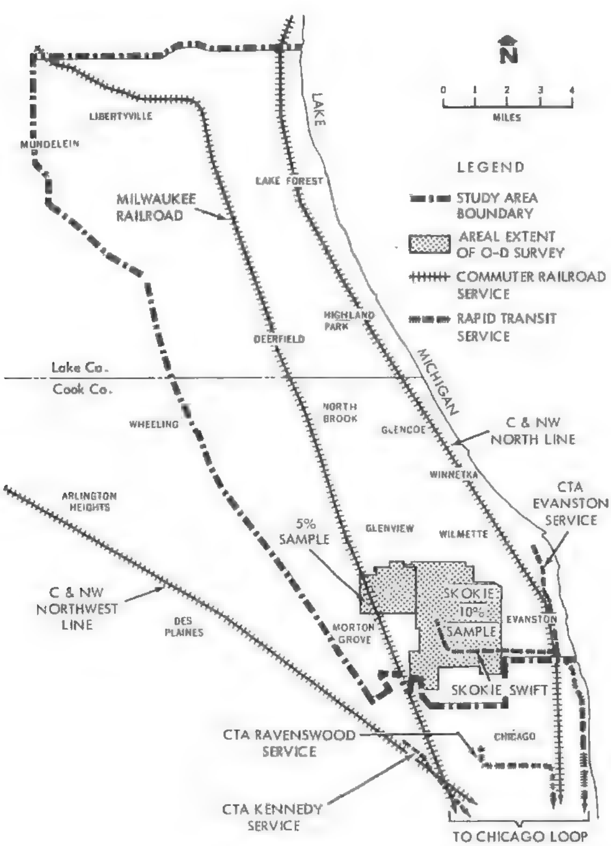


Figure 1. Corridor study area.

Analysis Needs

A high degree of accuracy in the ridership estimates for individual facilities was necessitated by study requirements. One element of the work program called for investigating modifications to the Skokie Swift CTA route, including an assessment of the effect these modifications would have on the parallel carriers. The analysis problem thus was not only to estimate transit ridership but also to allocate it to the correct submode with specific reference to passenger selection of certain competing rail routes.

Use of full regional networks and trip data was obviously not an appropriate strategy for this study. Although the North Shore suburbs form an urban area large in its own right, this same area represents only a small portion of the CATS coverage. Of the 681 CATS traffic analysis

zones, less than 15 percent are in the study corridor. Budgetary constraints dictated use of a travel analysis procedure that would not require involvement of the entire CATS system in the calculations. Conversely, the need for local detail favored subdivision of some study area traffic analysis zones.

As discussed in the next section, the study budget also ruled out taking surveys to obtain more travel data than were already available. Given these various considerations and constraints, the objective of the traffic analysis could be summarized as being to accurately estimate transit ridership in a corridor at a minimum cost.

Nature of Survey Data

Within the study area, recent survey data were available only for households within the village limits of Skokie, Illinois, and a portion of Morton Grove. These data had been collected in 1964 in connection with the HUD Skokie Swift and Intra-Skokie Mass Transportation Demonstration Projects. Compared to the study corridor as a whole, the home-interview survey area represented about 7 percent of the land area and a fifth of the population. Within the survey area, a 10 percent sample had been obtained within Skokie and a 5 percent sample within Morton Grove. Complete linked-trip data had been compiled by CATS. Supplemental to the home interview survey were transit rider postcard surveys that served to provide a larger sample for developing submodal choice and mode-of-arrival choice relationships.

The magnitude of survey area and population coverage was about what one might reasonably select given the opportunity to design a limited survey program for determination of study corridor travel patterns and modal-choice characteristics. However, the concentration of survey coverage into a contiguous area was a hindrance. Given the choice, one would obviously subdivide the survey coverage into a number of smaller unconnected areas designed to cover populations with more varied characteristics and travel-choice opportunities.

ANALYSIS STRATEGY

Corridor Analysis Development

The burden of carrying superfluous regional detail through the corridor analysis was removed by combining the traffic zones outside the study corridor and the Chicago CBD into large superdistricts. Concurrently, a number of large zones within the study area were subdivided. The resultant traffic zone system contained only 99 zones. Travel patterns were represented by trip tables compressed and reallocated to this modified zone system. Only trips internal to the study corridor or moving between the corridor and Chicago were retained. Suburban travel involving destinations external to the study corridor was deemed insignificant to transit considerations.

The highway and transit networks adapted for the study were prepared with fairly extensive detail in the corridor area and the Chicago business district. Skeletal network representations of the major facilities were used for the remainder of Chicago. Appropriate pseudo-links were used to represent minor road and feeder transit connections to the superdistrict centroids.

These actions reduced analysis costs markedly as compared to the typical transportation study. The ability to examine multiple plan alternatives in detail was correspondingly increased.

Choice Model Formulation

The choice-modeling strategy adopted was to use a chain of formulations, with each formulation describing the choice between 2 alternative mode categories. All formulations were developed in general accordance with the so-called utilitarian theory of travel mode choice (2). In brief, the pertinent elements of this working hypothesis are as follows:

1. Individual choice of mode is utilitarian. If an individual's unique perception of the travel disutility of each alternative could be measured for a given trip, his choice of mode could be absolutely predicted.

2. Description of individual utility perceptions with standard travel analysis techniques is affected by a multitude of chance errors. These, in sum total, can be described by the normal distribution error function.

3. The probability of free choice of a given travel mode is thus a probability density function of the disutility savings obtainable through use of that mode as compared to the alternate.

4. Deviations result from predictable influences such as captive riding and resistance to long trip length.

Application of the hypothesis makes use of the concept that a single measure can be used to describe travel disutility. The measure used in this study combines time, convenience, and dollar cost into a common unit of equivalent time called the utile. Mathematically, this utile is described as follows: Utile = A × (running time) + B × (walk and wait time) + C × (cost). The coefficients A, B, and C are unknown and must be determined as part of the calibration.

The formulations developed for binary mode selection applied to the following 4 levels of choice:

1. Prime modal split, automobile versus transit;
2. Submodal split, commuter railroad versus other transit;
3. Submodal split, 1 noncommuter transit mode versus another; and
4. Mode-of-arrival split, automobile arrival versus bus or walk arrival at a rail transit station.

Forecasting Model Chain

Study objectives made it desirable to split transit trips moving between the study area and Chicago into 4 submodes. These were commuter railroad, Skokie Swift, other rapid transit, and bus only. Because local trip rail use was minimal, it was decided to limit local trip model application to direct estimation of bus use.

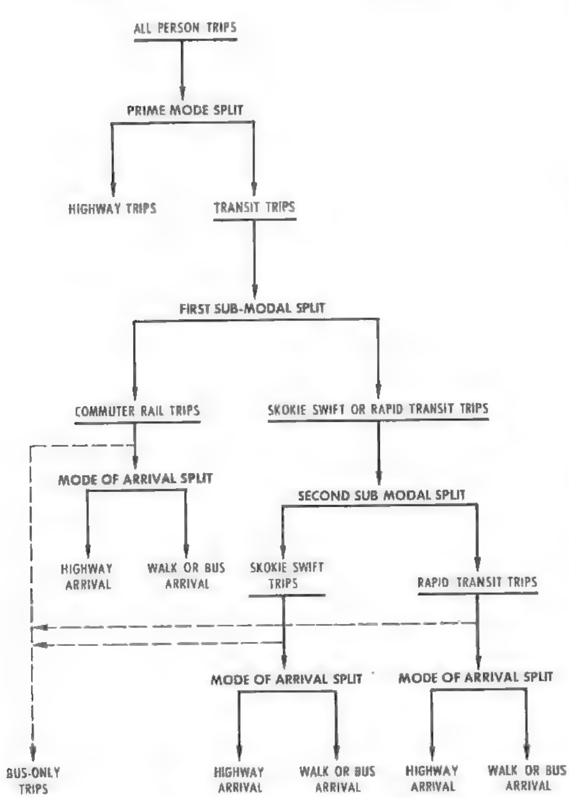


Figure 2. Modal split sequence for commuter travel.

The travel analysis was structured to start with a person trip travel estimate that had been developed from forecasts by CATS. The person-trip estimate was successively split on a trip interchange basis until allocation to the desired submode categories was complete. The sequence followed is shown in Figure 2. As a final adjustment, bus arrival trips were manually segregated from walk arrivals by using a transit station coverage measure.

The interchange impedance measures required by the modeling and forecasting strategy were provided by determining separately the minimum utility transit paths for each of the 3 private right-of-way travel modes. This was done for each private right-of-way mode by removing connections to the transit lines of the 2 modes not under consideration and then building the corresponding network description and minimum paths.

Automobile travel times, operating costs, parking charges, and terminal delays for the prime modal split were obtained in the normal manner.

For travel between the study corridor and Chicago, the "bus-only" trip estimate was accumulated as a residual. All bus lines were retained in each of the 3 transit networks, allowing bus-only paths for interchanges where the mode under consideration was inferior to bus. When trips were split to a path exhibiting no private right-of-way mode travel time, they were automatically assigned to the bus-only category.

Transit networks were coded by using the HUD transit planning programs. For choice-model calibration, the basic travel disutility components were accumulated separately for each of the various modes. However, once model calibration was complete, it was possible to convert all times and costs into utiles by using the equivalent time measure. This allowed the forecast networks to be coded in utiles instead of minutes and the inclusion of all times, convenience measures, and costs.

MODEL CALIBRATION TECHNIQUES

Prime Mode Choice

The prime mode-choice model required by the analysis strategy was calibrated on the basis of observed percentage of home-to-work trips made by transit by residents of the Skokie-Morton Grove survey area. Analysis was limited to interchanges connecting zones of the survey area to attractions in the survey area itself, in Chicago, and in the populous Evanston suburb to the east. The applicable portion of the compressed 99-zone system already described was used except for minor adjustments to match survey coding.

Calibration was facilitated by manually preparing a record for each zonal interchange containing all pertinent interchange information including the number of trips for each mode and the various time and cost measures describing the minimum paths for each mode. The resultant set of records contained all the data needed for model calibration. Computerized versions of this most useful technique have been well described in the literature (3).

Prior to regression analysis, the validity of the proposed formulation was checked by plotting observed percentage of transit versus an initial determination of the difference in disutility among alternative modes. Because at this point the coefficients needed to construct the utile measure had not been determined for the study area, it was necessary to make use of weights developed in previous applications (4). The plotted points did approximate a normal distribution function as anticipated, with the exception that most interchange data, where commuter railroad was the predominant transit mode, showed a higher transit use than would be expected. Commuter railroad travel parameters were subsequently adjusted, as is described later.

The next step was to determine the final weights for computation of the utile. This was done with regression analysis using a modification of the HUD transit planning regression program. Because the model structure is based on the normal distribution error function, application of linear regression required transformation of the dependent variable, percentage of transit. This was accomplished by relating each observed percentage of transit to a standard score, a technique that has been described by Bevis (5). A standard score is the cumulative normal distribution function expressed in values from -281 to +281. Zero percent transit use equates to -281, 50 percent equates to 0, and 100 percent equates to +281.

The first attempt at determining parameter weights used unrestrained regression analysis. This proved to produce mediocre results. A substitute procedure was developed involving use of several trial weight combinations to precompute the utile. Corresponding trial regressions were then run with the following formulation: Percent transit (standard score) = A + B × (transit utiles - highway utiles). Each resultant equation was next identified by its true percentage of error, computed as estimated trips divided by observed transit trips. The percentage of error values were then examined, and the choice was made of those weights that would produce 0 percent error.

The weights chosen for calculation of the utile were 1.0 for running time, 2.3 for walk and wait time, and 3.3 for cost in cents. A final regression equation was then determined for prime modal split as follows: Percent transit (standard score) = 29.2 - 2.76 (transit utiles - highway utiles).

The R statistical measure for this equation is 0.55, which describes the ability of the equation to reproduce the standard score of the observed percentage of transit. This particular R-value computation is not directly comparable to results from other modal-choice analyses.

The phenomena, already mentioned, of higher observed commuter railroad use than would be expected from the network measures appeared to result from 2 effects peculiar to that mode. One effect concerns the waiting time for the train as perceived by the commuter. Study area railroad service frequency is low compared to alternative modes, but schedule adherence is excellent. It was postulated that, for such relatively infrequent but rigidly scheduled transit services, the perceived wait may be less than the calculation used would indicate. Accordingly, the average wait computation for commuter railroad was reduced from a half of the headway to a quarter of the average headway.

The second effect concerned the commuter railroad ride itself. The study area commuter trains, in contrast to the rapid transit cars, have roomy seating, air conditioning, a minimum of center city stops and a seats-for-all policy. Comfort of the ride was taken into account by multiplying commuter railroad travel times by 0.7. These 2 modifications were used in both the modal choice and submodal choice analyses.

Submodal Choice

The augmented travel survey data used in the submodal choice calibration and pertaining primarily to route choice between the Skokie Swift and other rapid transit services had already been used in earlier investigations (6). In the present study, the survey observations were reapplied in accordance with the required modeling structure. The utile measure determined in the prime modal choice analysis was used to develop the following relationship: Percent Skokie Swift (standard score) = -14.68 × (Skokie Swift utiles - rapid transit utiles). The constant term was insignificant and the R-value for the equation was 0.92.

The matter of the constant term is of interest in that the theory behind the model formulations postulates that there should be no constant term for free-choice modal selection. That is to say, at 0 disutility difference, each mode should attract 50 percent of the total travel. In submodal split, captivity to any specific submode should be nonexistent. Thus, the lack of a significant constant term in the submodal regression equation bears out the theory.

In the development of the prime modal choice model, possible automobile and transit captivity (2) was not investigated. Accordingly, the resulting equation cannot be said to pertain exclusively to free-choice riders. For this reason, it is thought that the constant term in the prime modal split equation is caused by captivity effects, primarily the high-income worker captivity to the automobile. Further analysis would be required to ascertain the validity of this presumption.

The available survey data were not sufficient to allow formal calibration of a commuter railroad versus rapid transit submodal split formulation. As a substitute, the rapid transit route choice formula was applied after an adjustment in sensitivity. The adjustment was primarily keyed to terminal relationships in the Chicago business district. The formulation used was as follows: Percent commuter railroad (standard score) = -7.5 (commuter RR utiles - fastest rapid transit utiles).

The mode-of-arrival model was analyzed in much the same manner as the other choice models with the exception that a combined utile measure was not used. The equation for this model was as follows: Percent bus or walk (standard score) = 46.5 - 5.2 (difference of cost) - 4.2 (difference of walk and wait time).

MODEL EVALUATION TEST RESULTS

Three separate tests were conducted of the entire model chain to ensure that the travel forecasting tools being used were adequate for the needs of the study. Two tests involved using the forecasting sequence to prepare estimates of current conditions for comparison with ground counts or similar data. The third test examined the validity of the models under changing conditions. This was accomplished by producing an estimate of riding in the period immediately prior to opening of the Skokie Swift service. In all 3 tests, transit use was estimated for the entire survey corridor.

In the first test, an estimate was prepared of total 24-hour transit use for trips between the city of Chicago and the study corridor. A person trip estimate for 1970 and networks representing existing conditions were used. The resultant assigned volumes were checked against passenger counts. Figure 3 shows the results of a comparison made at Oakton Street, the most southern line in the study corridor where train counts would be free of the local Chicago riding, which was not estimated. Note that the counts were adjusted to remove through riding from beyond the study area.

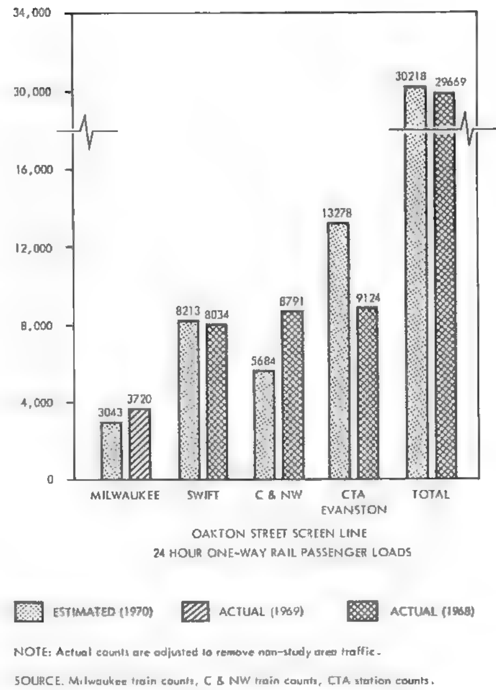


Figure 3. Comparison of synthetic estimate versus actual counts.

The assignment to the commuter rail lines was low, even with the model adjustments made as previously mentioned. On the other hand, both the total screenline forecast and the critical Skokie Swift forecast match counts within 2 percent.

In the second test, estimates of bus use were prepared for each of the suburban bus companies. This estimate was built up from 3 components: estimated bus-only travel between the study corridor and Chicago, estimated bus use for local travel, and estimated bus use for access to rail stations. This test involved all the models and techniques that had been developed. The standard of comparison was average weekday bus revenue passengers. Table 1 gives the results, which show an overall correspondence with 5 percent.

TABLE 1

COMPARISON OF SYNTHETIC BUS RIDER ESTIMATE VERSUS REPORTED PATRONAGE

Bus Company	Weekday Ridership	
	Actual	Synthetic Estimate
Evanston Bus Company	26,486	26,258
United Motor Coach Company	5,697	7,718
Glenview Bus Company	3,125	3,151
Total	35,308	37,127

For the third test, the Skokie Swift line was removed from the transit networks, and the model chain was rerun. The difference in results with and without the Skokie Swift service was then compared with prior mode use data obtained from Skokie Swift riders by postcard surveys in both 1964 and 1966. This test was extremely successful; the prior mode estimate produced by the models fell essentially between the results of the 2 surveys. The comparison is shown in Figure 4.

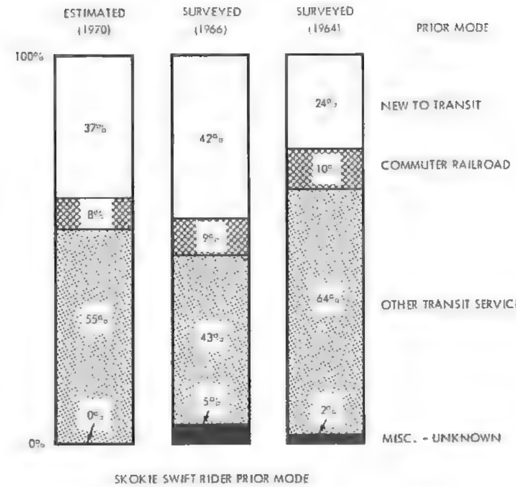


Figure 4. Comparison of synthetic estimate versus survey data.

CONCLUSION

In the example under discussion, use of corridor-level planning techniques and submodal analysis allowed stringent project requirements to be met. The objective of accurately estimating transit ridership at low cost was adequately accomplished with particular reference to estimating shifts in riding associated with modifications to the Skokie Swift. The cut-down networks and trip tables made possible by subarea analysis facilitated inexpensive testing of multiple alternatives.

The calibration of the models from data gathered in a subarea and the subsequent testing and use on the entire study corridor was done from necessity, but the results lent credence to the accuracy of the models.

The difficulties experienced in estimating commuter railroad use and the related understatement of the estimates provided an unexpected degree of statistical support for the importance of amenities such as comfort and schedule adherence. This might be a fruitful area to explore with more extensive survey data.

Perhaps the most significant aspect of the modeling work is that it was accomplished, as has already been discussed, in general conformance with postulated theory. Comparison of results with the hypotheses of the utilitarian theory of mode choice (2) are most encouraging, although it is obvious that a number of theoretical questions remain in need of further investigation. In any case, the techniques of the study described here do appear to make broader application of predictive modeling both feasible and useful.

REFERENCES

1. Pratt, R. H., and Bevis, H. W. Mass Transportation Study for the Village of Skokie as Agent for the North Suburban Transportation Council, 1969-71.
2. Pratt, R. H. A Utilitarian Theory of Travel Mode Choice. Highway Research Record 322, 1970, pp. 40-53.
3. St. Louis Mode Choice Model Development. Alan M. Voorhees and Associates, Inc., prepared for the Missouri State Highway Commission and the Illinois Division of Highways, Sept. 1969.
4. Shunk, G. A., and Bouchard, R. J. An Application of Marginal Utility in Travel Mode Choice. Highway Research Record 322, 1970, pp. 30-39.
5. Bevis, H. W. Estimating a Road-User Cost Function From Diversion Curve Data. Highway Research Record 100, 1965, pp. 47-54.
6. Pratt, R. H., and Deen, T. B. Estimation of Sub-Modal Split Within the Transit Mode. Highway Research Record 205, 1967, pp. 20-30.

DEMAND FOR TRAVEL ON THE CANADIAN AIRWAY SYSTEM

P. M. Pearson*, Parkin, Searle, Wilbee, and Rowland, Toronto, Canada

•A STATEMENT of future transportation needs is required so that transportation systems can be effectively and efficiently planned. Demand forecasts are fundamental planning variables and serve as a basis for generating long-term development plans, establishing investment priorities, and designing and implementing physical facilities.

However, many forecasts of demand, especially for air travel, have proved to be inaccurate. This has resulted in the provision of facilities that were incapable of serving actual traffic volumes. For example, the first aeroquay at Toronto International Airport, which was designed for 3 million passengers per year, was to provide adequate capacity until 1970 (1). However, in 1966, 2 years after the aeroquay was opened, there were almost 3-1/2 million passengers (2). On a larger scale, forecast and actual domestic air passenger-miles for the United States are shown in Figure 1.

The deficiencies of existing air travel demand models primarily result from the inability to quantify many of the underlying socioeconomic and transport system factors. For example, many models relate the total trips generated to total population. These models cannot account for changes in individual trip-making behavior. Other models do not explicitly include transport system factors and, therefore, cannot forecast traffic that will be generated because of technological advances. Furthermore, existing models examine only pairs of cities taken one at a time, and the competition of destination attractions cannot be taken into account.

The objective of this paper is to present an air travel demand model to overcome the preceding deficiencies. The modeling technique is based on systems theory and particularly on linear graph analysis. The technique is applied to business travel on the Canadian Domestic Airway System. The changes in traffic volumes as related to changes in cost and time of travel are derived for selected city pairs.

STATEMENT OF THE PROBLEM

The primary purpose of the linear graph model presented in this paper was to simulate the demand for intercity business travel on the Canadian Domestic Airway System. The problem may be stated as follows:

Given

1. An origin city consisting of people with specified incomes,
2. A travel network consisting of air links between the origin city and all destinations on the system, and
3. Destination cities consisting of various land use and employment types;

Find

1. The total number of annual business air trips originating at the origin city,
2. The assignment of these trips on the available travel links, and
3. The number of annual business trips from the origin city arriving at the destinations in the system.

*Mr. Pearson was with the Memorial University of Newfoundland in St. John's, Newfoundland, Canada, when this research was performed.

Sponsored by Committee on Transportation Forecasting.

MODEL FOR BUSINESS TRAVEL ON THE CANADIAN DOMESTIC AIRWAY SYSTEM

The system examined in this model is shown in Figure 2. It includes 11 major Canadian airport regions as follows:

1. Vancouver, Victoria, and New Westminister;
2. Edmonton and Calgary;
3. Regina and Saskatoon;
4. Winnipeg;
5. Toronto;
6. London and Windsor;
7. Ottawa;
8. Montreal;
9. Quebec City, Trois Rivières, Bagotville;
10. The Atlantic Provinces; and
11. Newfoundland.

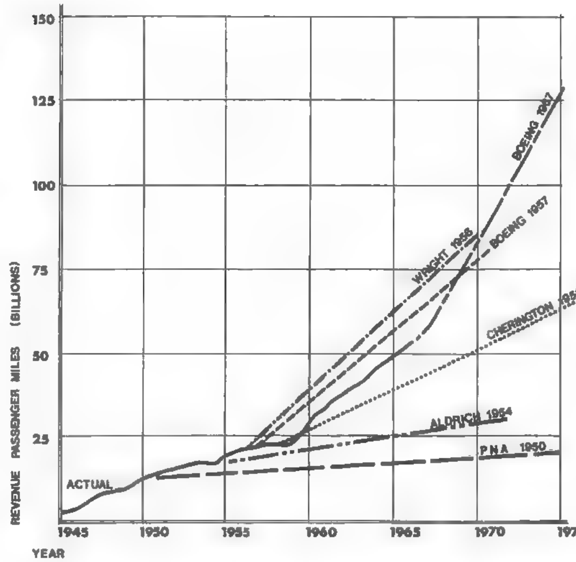


Figure 1. Comparison of forecast and actual air passenger-miles (U.S. domestic market, 3).

The components of the system included origin cities, destination cities, and all nonstop airway links between the origin and destination cities. The origin and destination cities are shown as links 001 and 011 in Figure 2. For any particular run of the model, 1 city acted as origin and all others as destinations. This allowed 1 line of an origin and destination table to be constructed per run. The origin or destination city included all major centers of population served by the airport. For example, link 006 is the London and Windsor region and included Woodstock, St. Thomas, Chatham, Sarnia, and Wallaceburg, as well as

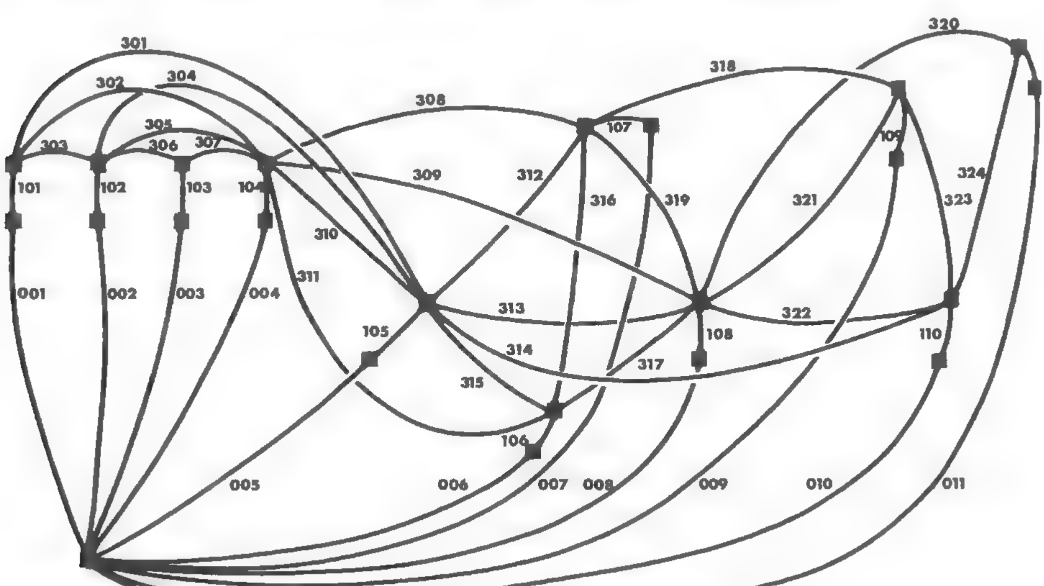


Figure 2. Systems graph for the Canadian domestic airway systems.

London and Windsor. Links 101 to 111 were measures of airport access and egress. Links 301 and 324 were the nonstop airway links between the cities.

Measurement on Components

Linear graph analysis requires that the following specifications be met:

1. The individual system components must be quantitatively describable by 2 fundamental variables. These variables are a flow variable y and a complimentary pressure variable x that causes flow.
2. The components are connected at their ends (vertexes) to yield a model for the entire system. The interconnected model must satisfy the 2 generalized Kirchoff laws. The first law states that the algebraic sum of all flows y at a vertex is zero. The second law states that the algebraic sum of all pressures around any closed loop of the system must be zero.
3. The flow and pressure variables must be related by a linear or nonlinear function.

The y-variable for the intercity air travel network is person trips per year. This satisfied the first Kirchoff law and eliminated the necessity of modeling for storage within the system. That is, all business travelers are assumed to return to their origin over the yearly period.

The x-variable is postulated to be a value measure used by the travelers in making an air trip. It is analogous to the portion of the travel potential of an origin area that is used up as a trip is made and thus is that pressure causing the flow.

The reasons for the preceding postulates are as follows:

1. If it is believed that the making of a trip can be simulated within a reasonable degree of accuracy, then it follows that there is some underlying process made by the traveler in making such a choice.
2. The traveler will act as a free agent and attempt to optimize his degree of satisfaction.
3. Relating reasons 1 and 2 to a value measurement used in travel allows the origin pressure to dissipate as the trip is made, thus satisfying the second Kirchoff law.

Terminal Equations of Components

Linear graph analysis requires that each link be individually described with a relationship between the flow and the pressure variables. This relationship is called the terminal equation of the component. For the airway model, 3 forms of terminal equations were required, one for each of the component types—destination, airway link, and origin.

Destination Area Components

The terminal equations of the destination cities can be expressed as

$$Y_k = A_k X_k \tag{1}$$

where

Y_k = number of business trips per year arriving at destination k,
A_k = attraction of city k, and
X_k = remainder of the perceived cost of travel that is used up across city k.

The attraction variable A_k for business trips was postulated to be a function of employment of the form

$$A_k = f\ (\text{employment}) \tag{2}$$

A recent study by Air Canada (4) provided information on the trip-making characteristics of various employment types. The study provided employment trip-making characteristics, and the following relationship was developed:

$$A = \phi \left(B_s \frac{e_{sk}}{e_{s,ave}} + B_H \frac{e_{Hk}}{e_{H,ave}} + B_L \frac{e_{Lk}}{e_{L,ave}} \right) \quad (3)$$

where

A_k = relative attraction of a destination;
ϕ = calibration constant;
e_{sk}, e_{Hk}, e_{Lk} = employees of a destination city in the service, heavy, and light industries respectively;
$e_{s,ave}, e_{H,ave}, e_{L,ave}$ = number of employees in the service, heavy, and light industries respectively in the average city of the network;
B_s, B_H, B_L = trip attraction characteristics of each employment type (these were found to be 0.452, 0.362, and 0.185 respectively).

The employment factors ($e_{ik}/e_{i,ave}$) for the destination cities were derived from data of Dominion Bureau of Statistics (6).

The total number of employees in service, heavy, and light industries was calculated for each city. These were then divided by the average number of employees in the service, heavy, and light industries in the 11 cities in the system. These dimensionless numbers were then multiplied by the appropriate constant, and the attractions were derived. The attraction values are given in Table 1.

Air Link Components

The terminal equations of the air links were postulated to be of the form

$$Xij = R(y) Yij \quad (4)$$

where

Xij = perceived value or cost used up by the business traveler in crossing the link,
Yij = flow in person per year across a link ij, and
R(y) = resistance to flow.

It is hypothesized that the resistance function can be expressed as

$$R(y) = C(y) + T(y) \quad (5)$$

TABLE 1

DERIVED ATTRACTIONS FOR THE 11 AIRPORT REGIONS

Airport Region Code	Name	$0.452 \frac{e_s}{e_{s,ave}}$	$0.363 \frac{e_H}{e_{H,ave}}$	$0.185 \frac{e_L}{e_{L,ave}}$	Total Attraction
001	Vancouver	0.4342	0.1174	0.3599	0.9915
002	Edmonton	0.3838	0.1209	0.3179	0.8226
003	Saskatoon	0.1353	0.0566	0.0421	0.2340
004	Winnipeg	0.2454	0.1107	0.0600	0.4161
005	Toronto	1.4825	1.2356	0.2263	2.9444
006	London	0.2338	0.3361	0.0462	0.6161
007	Ottawa	0.1870	0.0157	0.1083	0.3110
008	Montreal	1.2101	1.6705	0.0462	3.0281
009	Quebec	0.1764	0.1525	0.0864	0.4153
010	Maritimes	0.4282	0.1624	0.4869	1.0775
011	Newfoundland	0.1001	0.0130	0.1190	0.2321

where

$C(y)$ = function of cost to cross a link in units of cents per person per mile; and
$T(y)$ = function of time to cross a link in units of minutes per person.

The functions $C(y)$ and $T(y)$ can incorporate the capacity and scheduling parameters of public carriers.

The cost function is of the form

$$C(y) = M\frac{c}{m} + \frac{c(q)}{m} \tag{6}$$

where

$C(y)$ = cost in cents per passenger per mile;

$\frac{c}{m}$ = cost in cents of the air fare per mile on a link;

$\frac{c(q)}{m}$ = quality of travel costs in cents for a link-mile; and

M = length of the link in miles.

It was assumed that travelers perceived cost directly. The cost term contains no model-calibration term. The quality of travel, $c(q)/m$, reflects the comfort and convenience of traveling by the air mode as compared to another mode. This cost at the present time is difficult to measure.

The time to cross a mile of travel link is given by the formula

$$t(y) = KM\frac{t}{m} + \frac{t(d)}{m} \tag{7}$$

where

$\frac{t}{m}$ = time in minutes per passenger to cross a link;

$\frac{t(d)}{m}$ = time delay in minutes per mile of travel link;

K = constant defining of the perceived travel time costs in cents per minute; and

M = length of the link in miles.

The term $M[t/m + t(d)/m]$ can be approximated with the mean journey time concept developed by Morlok (7).

A capacity constraint can be imposed on public carriers. The constraint states that, if available capacity is exceeded, no additional trips can occur on a particular link. The constraint can be stated in terms of available space as

$$ES = B_1 S_1 + B_2 S_2 + \ldots \tag{8}$$

where

ES = effective seats available; and
B_1 = proportion of available seats S_1 that are demanded in a given time period of the day.

The total resistance on a travel link is given by

$$\begin{aligned} R(y) &= C(y) + T(y) \\ &= M\left[\frac{c}{m} + \frac{c(q)}{m}\right] + KM\left[\frac{t}{m} + \frac{t(d)}{m}\right] \end{aligned} \tag{9}$$

or

$$R(y) = \infty \quad \text{for } y_{ij} > ES$$

where the terms are as previously defined. No capacity constraint problems were encountered in the networks examined in the project.

The airway link resistances were thus calculated by considering air fares, route travel times, and scheduled departure frequency. The values for these resistances are given in Table 2.

Airport Access Link Resistances

The airport access and egress link resistances (101 to 111) for business air passengers were calculated by considering travel costs, travel times, passenger processing times, and passenger insurance time (8). The resistances were calculated from Eq. 14 with K = 10.0 per minute. The choice of value of K is discussed in a following section. The access and egress resistance values are given in Table 3.

For destination airports within 300 miles of the origin area, it was necessary to modify the resistances of the airport egress links. Within this trip length range, there is considerable competition among air and other modes of travel. It was felt that some measure of competition should be included. Furthermore, this measure was included in the access links so that other traffic using the same airway link for longer trips would not be similarly penalized. The competition measure was a scalar that was multiplied by the access resistances. Thus, for the previously described egress links,

TABLE 2

RESISTANCES FOR CANADIAN DOMESTIC AIRWAY NETWORK

Link[a]	Time (min)	Cost (dollars)	$K_2 = 5$	$K_2 = 10$	$K_2 = 20$
301	972	109	15,760	20,620	30,340
302	696	63	9,780	13,060	20,220
303	270	32	4,550	5,900	8,600
304	870	89	13,250	17,600	26,300
305	606	43	7,330	10,360	16,420
306	408	29	4,940	6,980	11,060
307	294	26	4,070	5,540	8,480
308	876	63	10,680	15,060	23,820
309	876	63	10,680	15,060	23,820
310	408	52	7,240	9,280	13,360
311	912	50	9,560	14,120	23,240
312	144	19	2,620	3,340	4,780
313	138	23	2,990	3,680	5,090
314	834	48	8,970	13,140	21,480
315	246	15	2,730	3,960	6,420
316	798	28	6,790	10,780	18,760
317	852	40	8,260	12,520	21,040
318	540	22	4,900	7,600	13,000
319	132	11	1,760	2,420	3,740
320	864	60	10,320	14,640	23,280
321	192	13	2,260	3,220	4,140
322	408	28	4,890	6,930	1,010
323	792	22	6,160	10,120	18,040
324	864	37	7,020	12,340	20,980

[a]Figure 2 shows relative position of links.

TABLE 3

ACCESS AND EGRESS RESISTANCE VALUES
FOR K = 10.0 CENTS PER MIN

Link	Access and Egress	Egress for < 300 Miles	Origin
101	1,510		
102	980		
103	1,020		
104	1,040		
105	1,115	13,500	London
106	1,320	13,500	Toronto
107	880	14,300	Montreal
108	1,160	14,300	Ottawa
109	1,710	12,000	Montreal
110	1,940		
111	1,940		

$$R_{egress} = SR_{link} \tag{10}$$

where

R_{egress} = resistance of the short-haul airport egress link (values are given in Table 3);

S = scalar that is postulated to account for the trade-offs made by business travelers in choosing air for a trip shorter than 300 miles (cost and time parameters were derived from Norhling 9); and

R_{link} = egress link resistance derived from Eq. 14.

Origin Area Components

The origin area can be characterized as a known flow driver of the form

$$Y_i = y_s \tag{11}$$

where

Y_i = flow from origin i in annual business trips; and

y_s = specified flow value taken from actual data.

The relationship is necessary in calibrating the model for a given network. The flow values were taken as the total flows originating in a city for destinations on the network.

As it will be shown, the origin areas are modeled as an individual link of the system. Thus, associated with the origin area flow is its complementary pressure variable. The magnitude of the pressure variable can be derived from the solution of the linear graph network. This pressure is related to the origin travel volumes because it is that pressure that was necessary to create the flows.

Therefore, the pressure variable of the origin area is postulated to be the travel potential that created the traffic volumes. As will be shown, the travel potential can be characterized by the following equation

$$X = A(IP) + B \tag{12}$$

where

X = travel potential in cost per year (which is the perceived cost used up as the trip is made);

I = annual average income of the origin area population P;
P = origin area population; and
A, B = constants (A implicitly describes the travel attributes required to make a business trip, and B implicitly describes the threshold value attached to the travel attributes before any trip can be made).

Systems Graph

The systems graph is a set of terminal graphs (i.e., a graphical representation of a component and its terminal equation) connected at the vertexes to form a one-to-one correspondence with the components of a physical system. Figure 2 shows a systems graph for the Canadian Domestic Airway System.

Systems Equations

To construct the air travel demand model required the derivation of both the chord and branch formulation equations of the system. Both of these formulation methods can be found elsewhere (11, 12, 13).

The graph procedure is illustrated by the following generalized results of the formulation techniques. From Figure 2, with city 001 as a known flow driver, the independent chord formulation equations can be written as

$$\begin{array}{ccc} BRB^T Y_c & + 0X_{c1} & = 0 \\ Y_{c1} & U & \end{array} \tag{13}$$

where

B = matrix whose coefficients represent the manner in which the systems is connected (i.e., the coefficients of the cut-set equations);
B^T = transpose of B;
R = diagonal matrix (whose entries include the link resistances and the destination city attractions);
Y_{c1} = known flow value for city 001;
Y_c = column matrix whose entries are the unknown chord flows; and
X_{c1} = unknown pressure or travel potential for city 001.

Removing the last equation from the set makes it possible to derive the unknown chord flows. Substitution into the cut-set equations yields the branch flows. The unknown pressures are found by substitution into the terminal equations. The origin area travel potential can then be solved from the last equation of the set. The procedure is repeated for each city.

With the origin area travel potentials available, branch formulation equations can be derived. The branch equations with city 001 as the origin is given by

$$\begin{array}{ccc} UY_{c1} & + ARA^T X_{B1} & = 0 \\ 0 & XB & \end{array} \tag{14}$$

This system of equations can be solved for all pressures and flows. It is the branch model that is capable of expressing the generations of air trips in relation to the system parameters.

CALIBRATION OF THE MODEL

The model was calibrated to obtain a value of K for the resistance function. Three cities were chosen at random—one to represent a large city, one to represent a medium city, and one to represent a small city. The city names were grouped according to originating traffic volumes as follows: large volumes, Toronto and Montreal; medium volumes, Vancouver, Edmonton and Calgary, Ottawa, and Winnipeg; small volumes, Atlantic Provinces, Saskatoon and Regina, London and Windsor, Quebec, and Newfoundland. One name was then selected from each group: Montreal (large), Vancouver (medium), and London and Windsor (small).

TABLE 4

ONE-WAY AIR PASSENGER BUSINESS TRIPS IN 1964

Origin[a]	Destination[a]											Total
	001	002	003	004	005	006	007	008	009	010	011	
001	x	30,137	6,080	11,897	12,053	3,238	224	5,218	1,082	1,326	511	73,766
002	18,932	x	10,784	10,277	7,034	2,227	1,809	3,848	457	1,160	339	56,867
003	2,019	6,656	x	4,524	1,852	740	524	1,123	257	222	122	18,039
004	5,540	7,796	7,147	x	7,532	4,740	2,843	5,051	1,284	1,541	479	43,953
005	25,570	26,127	11,338	39,043	x	19,667	52,310	97,108	24,106	36,127	9,293	340,689
006	866	1,276	625	2,200	1,094	x	2,412	4,838	1,341	1,281	468	16,421
007	1,439	1,936	1,249	3,872	15,868	5,615	x	2,670	6,853	3,779	1,712	44,993
008	7,352	8,485	5,050	20,204	71,300	30,355	10,043	x	17,426	42,188	16,244	228,647
009	253	208	142	589	2,414	824	1,652	4,298	x	1,936	486	12,796
010	397	497	252	692	4,029	1,284	1,475	6,083	3,181	x	3,038	20,928
011	166	131	108	276	1,135	291	536	3,139	771	3,273	x	9,826

Source: Canada Yearbook, 1967 (18).

[a]Names of origins and destinations identified by 3-digit code are given in Table 1.

The chord formulation equations were then derived for each of these cities. The 1964 business air traffic volumes (Table 4) were used as known flow drivers. The values of the flow drivers are given in the right column of Table 4. The chord equations were solved for values of K = 1.0, 5.0, 7.5, 10.0, 15.0, and 20.0 for each of the 3 cities (with 1 city acting as origin for each run).

The attraction calibration constant ϕ was set equal to 0.0001 to avoid numerical problems with matrix inversion. The actual and predicted distribution results were statistically compared.

The results of the statistical tests for the 3 cities with K = 5.0, 10.0, and 20.0 cents per mile are given in Table 5. The results show that the statistically best agreement

TABLE 5

COMPARISON OF DISTRIBUTION ERRORS FROM MODEL CALIBRATION

Distribution	Percentage Errors		
	K = 5.0	K = 10.0	K = 20.0
From Vancouver			
Edmonton	-9.55	-0.99	9.98
Saskatoon	3.04	-0.18	-10.51
Winnipeg	-0.01	6.99	12.40
Toronto	-8.01	-4.16	-4.34
London	79.28	6.58	-16.74
Ottawa	10.39	6.88	-9.44
Montreal	-0.36	-9.24	-29.95
Quebec	34.29	10.72	-23.84
Maritimes	32.50	5.28	-31.75
Newfoundland	14.87	-18.59	-55.38
From London			
Vancouver	22.40	11.32	-5.20
Edmonton	3.84	-6.82	-21.87
Saskatoon	5.12	-3.20	-16.96
Winnipeg	-2.32	3.55	11.59
Toronto	-16.27	-1.74	23.40
Ottawa	-5.39	3.19	15.88
Montreal	-3.06	-0.02	2.34
Quebec	-1.94	-8.57	-16.23
Maritimes	17.64	4.84	-12.88
Newfoundland	10.47	-11.32	-36.11
From Montreal			
Vancouver	25.34	2.80	25.44
Edmonton	37.98	11.74	-19.92
Saskatoon	22.06	3.05	-21.86
Winnipeg	1.82	-0.91	-5.38
Toronto	-4.27	1.01	7.04
London	4.44	0.95	-6.40
Ottawa	-2.32	12.29	36.80
Quebec	-21.92	-9.10	17.20
Maritimes	-3.76	-4.75	-7.55
Newfoundland	4.71	1.06	-5.18
Average percent of error	13.61	5.71	17.32
Root mean square error	21.02	7.18	21.54
Algebraic sum of errors	2.48	0.04	1.95

was obtained for K = 10.0 cents per minute. This represents a value of $6 per hour, which closely approximates the value of time for business air travelers established by the Port of New York Authority (14) and Simpson (10).

RESULTS OF THE CHORD FORMULATION OF THE MODEL

Chord formulation equations were constructed for the entire system. The attraction values of the destination components (Table 1) were multiplied by 0 = 0.001. The resistance values for the airport access links (Table 3) and the airway links (Table 2) used in the model had a value of K = 10.0 cents per minute. The values of the known flow drivers are given in the total column of Table 4. These figures are the actual 1964 origin and destination data for business trips on the airway system.

The model was run with 1 city, in turn, acting as origin and all others acting as destinations. The results are given in Table 6. The statistical tests for the total airway model are given in Table 7. The average percentage of error was 7.90, the root mean square error was 11.01, and the algebraic sum of the errors was 0.18. In light of existing airline system load factors reported at 60 to 65 percent (15), the model results are well within acceptable limits for planning purposes.

Table 8 gives the airway link assignments, which were calculated simultaneously with the distribution results. Unfortunately, no data were available to verify the assignment results. They are included in the paper for completeness of model results.

DERIVATION OF ORIGIN TRAVEL POTENTIAL FUNCTION

The origin area travel potentials, which were derived from the solution of the chord equations, are given in Table 9. Table 9 also gives the average weekly salary, population, and the product of salary and population for each origin city.

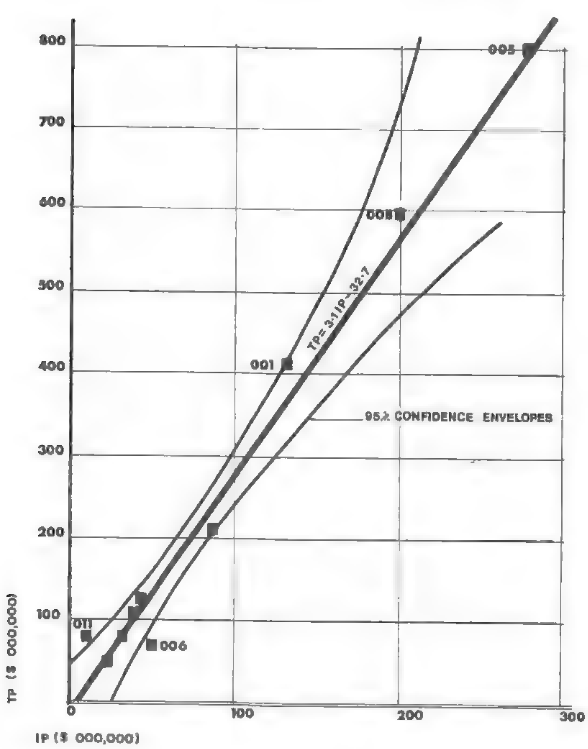


Figure 3. Comparison of data with travel potential regression function.

Travel potential was used as a dependent variable to perform a linear regression analysis on the data. The travel potential function was assumed to be the form of Eq. 17, namely,

$$TP \doteq A(IP) + B \qquad (15)$$

The equation was found to be

$$TP = 3.1(IP0 - 32.7) \qquad (16)$$

for $16.8 < IP < 281.0$.

The coefficient of correlation squared was 0.956. This indicates that 95.0 percent of the variation is the travel potential as explained by Eq. 16. The experimental values of travel potential and the product of weekly income and population are compared with the regression line shown in Figure 3.

The value of A(3.1) not only accounts for the money spent on travel but also reflects the other travel attributes of the population (i.e., time, and desire). Furthermore, if the average weekly income is a measure of the economic activity increases, it can be anticipated that business travel will also

TABLE 6

MODEL RESULTS FOR K = 10.0

Origin	Destination											Total
	001	002	003	004	005	006	007	008	009	010	011	
001	x	29,838	6,069	12,729	11,552	3,451	2,377	4,736	1,198	1,396	416	73,762
002	17,571	x	10,878	10,624	7,943	2,570	1,755	3,500	522	1,000	304	56,867
003	2,042	6,218	x	4,954	1,709	833	550	1,099	273	271	90	18,039
004	5,697	8,075	6,588	x	7,672	4,299	2,814	5,624	1,392	1,331	455	43,947
005	24,669	28,966	11,298	38,724	x	19,079	49,957	98,928	25,516	34,230	9,314	340,682
006	964	1,189	605	2,278	1,075	x	2,489	4,837	1,226	1,343	415	16,421
007	1,531	1,920	1,053	4,051	14,608	6,352	x	2,346	6,631	4,741	1,781	44,993
008	7,558	9,481	5,204	20,021	72,022	30,644	11,278	x	15,839	40,182	16,416	228,645
009	202	252	136	524	1,964	813	1,583	4,726	x	2,091	499	12,790
010	351	426	203	747	4,021	1,209	1,530	6,327	3,113	x	3,006	20,933
011	114	141	73	277	1,170	432	577	2,967	808	3,264	x	9,823

TABLE 7

RESULTS OF STATISTICAL TESTS ON AIRWAY MODEL

Origin	Destination										
	001	002	003	004	005	006	007	008	009	010	011
001	x	-0.99	-0.18	6.99	-4.16	6.58	6.88	-9.24	10.72	5.28	-18.59
002	-7.19	x	6.87	3.38	12.92	15.40	-2.98	09.04	14.22	-13.79	-10.32
003	1.14	-6.58	x	9.50	-7.72	12.57	4.96	-2.14	6.23	22.07	-26.23
004	2.83	3.58	-7.82	x	1.86	-9.30	-1.02	11.34	8.41	-13.63	-5.01
005	-3.52	10.87	-0.35	-0.82	x	-2.99	-4.50	1.87	5.85	-5.25	0.23
006	11.32	-6.82	-3.20	3.55	-1.74	x	3.19	-0.02	-8.57	4.84	-11.32
007	6.39	-0.83	-15.69	4.62	-7.94	13.13	x	-12.13	-3.24	25.46	4.03
008	2.80	11.74	3.05	-0.91	1.01	0.95	12.29	x	-9.10	-4.75	1.06
009	-20.16	21.15	-4.22	-11.04	-18.64	-1.33	-4.18	9.96	x	8.01	2.67
010	-11.59	-14.29	-19.44	7.95	-0.20	-5.84	3.73	4.01	-2.14	x	-1.05
011	-31.33	7.63	-32.41	0.36	3.08	48.45	7.65	-5.48	4.70	-0.27	x

Note: Average percentage of error = 7.90; root mean square error = 11.01; and algebraic sum of errors = -0.18.

TABLE 8

AIRWAY LINK ASSIGNMENTS TOTAL ONE-WAY PASSENGER TRIPS FOR MODEL K = 10.0

Link	Origin											Total
	001	002	003	004	005	006	007	008	009	010	011	
301	13,787	1,651	136	175	20,506	600	838	4,130	113	239	68	92,243
302	20,050	984	763	5,374	4,237	346	652	3,224	83	106	42	35,861
303	39,928	20,207	1,415	498	74	18	41	203	5	4	2	62,395
304	2,768	8,709	634	373	24,051	696	968	4,771	130	279	79	43,458
305	2,536	12,749	156	6,491	5,384	426	798	3,948	102	131	52	32,773
306	4,784	15,364	8,113	2,456	544	84	194	964	24	19	11	32,557
307	1,284	4,484	9,929	9,044	10,753	689	1,248	6,169	161	222	84	44,067
308	1,959	1,697	864	4,877	5,656	210	3,916	9,017	323	263	100	28,882
309	2,137	1,828	905	5,086	1,860	12	1,449	17,691	360	472	236	32,036
310	2,417	2,284	1,377	7,955	39,606	846	944	4,640	133	382	91	60,675
311	2,059	1,784	908	5,122	11,976	2,672	439	2,015	54	88	29	27,146
312	2,117	1,307	70	107	84,540	1,402	15,038	27,769	1,085	123	198	133,756
313	2,649	1,723	230	756	92,265	2,084	3,551	60,709	1,136	970	738	166,811
314	975	661	130	559	29,603	839	585	773	237	4,407	581	39,350
315	1,679	1,008	8	376	50,110	7,545	646	3,689	117	578	108	65,864
316	39	34	18	104	7,785	3,206	4,896	9,958	379	250	101	26,770
317	247	187	65	341	11,269	2,999	1,248	19,010	371	468	251	36,456
318	631	467	148	761	13,088	676	5,811	4,033	3,142	863	126	29,746
319	1,106	816	254	1,299	23,624	1,233	15,354	53,992	229	1,304	850	100,061
320	304	226	73	377	6,093	314	1,584	15,411	267	639	4,675	29,963
321	659	490	159	821	13,137	669	2,177	31,059	7,244	1,056	340	57,811
322	441	339	124	659	7,138	485	2,995	30,771	152	7,841	709	51,654
323	92	76	34	190	708	119	1,358	11,189	2,409	5,033	594	21,802
324	112	77	17	77	3,220	100	197	1,004	232	3,645	5,148	13,829

TABLE 9

DATA FOR ORIGIN AREA TRAVEL POTENTIAL FUNCTION

Origin	Derived Travel Potential (millions)	Origin Area Travel Volumes (thousands)	Origin Area Average 1964 Weekly Salary[a]	Origin Area 1964 Population[b] (millions)	Product of income and Population (millions)
001	413.098	73.766	90.47	1.465	133.1
002	220.252	56.867	83.48	1.069	87.0
003	89.244	18.039	74.90	0.516	38.6
004	130.502	43.953	76.28	0.615	47.0
005	866.723	340.689	92.82	3.053	281.0
006	66.457	16.421	83.09	0.650	53.9
007	111.924	44.993	80.72	0.646	52.1
008	597.808	228.647	85.89	2.328	200.0
009	52.253	12.796	74.08	0.403	29.8
010	104.328	20.928	73.21	0.600	43.9
011	92.421	9.826	66.72	0.253	16.8

[a]See (18).
[b]See (19).

increase. This fact is more or less substantiated by the regression equation. The population term merely reflects the number of people available to make trips.

There are 2 points outside of the 95 percent confidence level envelopes and these are Newfoundland (011) and London and Windsor (006). Because of its physical location, Newfoundland would require a greater travel potential to produce the same amount of trips as other areas. The travel resistance to make a trip is higher because the trips are longer. London and Windsor are in close approximation to heavily populated areas of New York State and Michigan. It can be concluded that a disproportionate number of air trips from this area are attracted to the United States. Therefore, the travel potential suggested by the regression function would create more trips on the airway system than were actually made.

APPLICATION OF THE BRANCH FORMULATION MODELS

Air Business Travel Demands

Branch formulation models were constructed for each origin area on the Canadian Domestic Airway System. The origin areas were modeled as pressure drivers, as discussed previously. The values of the pressure drivers were taken as the traffic potentials derived from the chord formulation models (Table 9). The branch formulation models permit demand elasticities to be calculated and demand curves for any particular origin and destination pair to be derived.

Income Elasticities

Sensitivity tests on the travel potential (and the total annual air business trips) generated by the origin areas were conducted with respect to the income variable. The income variable was changed for a number of selected cities. The resultant changes in travel potential and traffic volumes are given in Table 10. The income elasticities of Table 10 were calculated from

$$TP, I = \frac{TP_1 - TP_2}{I_1 - I_2} \frac{I_1}{TP_1} \tag{17}$$

where

TP, I = elasticity of travel potential or travel volume with respect to the income variable (the reported elasticities are for business air trips);
TP_1, TP_2 = original and final values of travel potential (flows);
I_1, I_2 = origin and final values of the average weekly salary.

In each case, the income term was increased by 10 percent. The resultant elasticities of travel volumes (Table 10) range from 2.01 for the Maritime Provinces to 1.00 for the Toronto airport area. The interpretation of the variation in the elasticities of income is as follows:

1. The lower travel potentials are associated with areas of relatively low incomes and population. An increase of 10 percent in average income would result from a large increase in economic activity. Therefore, the effect on business air travel volumes would also be large.
2. In an area of high economic activity such as Toronto, an increase of 10 percent in average salary may not reflect as great an increase in business transaction. Therefore, the increase in travel would be relatively inelastic.

Elasticity of Air Travel Costs and Times

The cost and time elasticities of air business travel were derived from the branch formulation models. The cost and time elasticities were derived for the Toronto and Montreal city pair. The origin area travel potentials were used as pressure drivers, and the air fare for link 313 (air route between Toronto and Montreal, Fig. 2) was increased and decreased by 5 and 10 percent with travel time remaining constant. The procedure was then reversed and the air fare was fixed while the travel time was changed ±5 and ±10 percent.

The resultant changes in travel volumes are given in Table 11. As would be anticipated, a decrease in air fare produces an increase in travel volumes. The anticipated increase is 5,100 annual yearly passengers for a decrease of 10 percent in air fare. The calculation of elasticity of business air trips produces a value of $A, p = -0.30$. Therefore, business air travel is inelastic by this model. This is in accordance with studies by a number of authors (14, 16) that business air travel is inelastic.

Business air travel is also inelastic with respect to travel time. The change in volume is 4,100 annual air business passengers for a change of 10 percent in the travel time. The change in travel time could correspond to a change in departure schedules as well as decrease in actual running time. The value of the elasticity is $A, T = 0.24$ and again is inelastic. One author (16) concluded that the time elasticity was greater than the cost elasticity. However, a 10 percent decrease in travel time amounts to

TABLE 10

RESULTS OF SENSITIVITY TESTS ON INCOME

City	Original Income	Final Income	Travel Potential[a]		Travel Volumes[b]		Elasticity of Traffic Volumes
			Original[c]	Final[d]	Original	Final	
Atlantic	73.21	80.53	104.3	125.0	20.9	25.1	2.01
Winnipeg	76.28	83.91	130.5	140.0	44.0	53.8	1.85
Vancouver	90.47	99.52	375.1	419.8	73.8	82.6	1.20
Montreal	85.89	94.48	595.8	658.1	228.6	255.0	1.05
Toronto	92.82	102.10	860.7	944.7	340.7	374.8	1.00

[a]In millions of cost units.
[b]In thousands of annual trips.
[c]From regression relationship.
[d]Original value plus 10 percent.

TABLE 11

DERIVED CHANGES IN TOTAL ANNUAL ONE-WAY BUSINESS PASSENGER VOLUMES FOR THE TORONTO-MONTREAL AIR ROUTE

Change	Cost		Time	
	Volume Change[a]	Elasticity	Volume Change[a]	Elasticity
-0.10	+5,100	-0.30	+4,100	-0.24
-0.05	+2,550	-0.30	+2,050	-0.24
0.00	—	—	—	—
+0.05	-2,550	-0.30	-2,050	-0.24
+0.10	-5,100	-0.30	-4,100	-0.24

[a]The volume changes are the total of generated plus diverted traffic.

approximately 14 min on link 313. With a value of K = 10 cents per minute, the 14 min represents $1.40. However, a 10 percent decrease in air fare is equal to a savings of $2.30. Therefore, in consideration of the preceding cost factors, the elasticity measurement of the branch model appears valid.

The elasticities incorporated in the model result not from the resistance change of the particular link but from the change of equivalent resistance for the origin, access and egress, and the destination links. The elasticities are therefore more or less constant for a city pair. With large changes of cost and time, the model may be in error. The assumed resistance functions may deviate considerably from the real world in the area of large cost-of-time changes. However, the range of the sensitivity measures should be adequate for most planning purposes.

CONCLUSIONS

The travel demand simulation procedure described by the chord formulation does not differ significantly from existing models with respect to the following:

1. There are coefficients of the model that must be calibrated;
2. The calibration constants are assumed to remain constant over the planning horizon; and
3. The origin flow values (or travel potentials) for the planning horizon must be estimated through some type of regression formulation.

However, the chord equation model does offer some advantages and these are as follows:

1. Generation, distribution, and assignment are considered as interdependent and are completed simultaneously for each origin;
2. The model considers the competitive attractions of all destinations on the system with respect to each origin;
3. The travel links are mathematically described by their time and cost parameters;
4. The model quantifies the interconnections of many of the variables relating to demand, and the variable and interconnection measurements are achieved at the aggregate level; and
5. The model is analytic and, therefore, no interactions or balancing procedures are required to determine the travel volumes on each link.

The nature of the sensitivity attributes of the branch formulation model offer some important features to the transportation planner, including the following:

1. Elasticity measures for system or origin attributes can be derived; and

2. The generation of trips due to changes in the system attributes can be identified.

ACKNOWLEDGMENTS

The research reported in this paper was sponsored by the Canada Department of Transport. Many of the data were provided by Air Canada, Canadian National Railways, and the Department of Highways, Ontario.

REFERENCES

1. Private Communication with A. R. Conboy, Chief Economist of Transportation Policy and Research Branch, Canada Department of Transport.
2. Forecasts of Aviation Activity. Canada Department of Transport, Ottawa, 1968.
3. Le Fevre, W. L. Changes in Growth Rates of Domestic Passenger Air Transportation. Papers A.T.RF., 1962.
4. Vancouver Study. Air Canada, Montreal, 1966.
5. Market Analysis Studies. Air Canada, Seasonal and Annual Reports, Montreal, 1959 to 1965.
6. Labour Force by Industry. In Census of Canada Vol. 3, Part 2, Queen's Printer, Ottawa, 1964.
7. Morlok, E. An Analysis of Transport Technology. Northwestern Univ., PhD Dissertation, 1967.
8. Hutchinson, B. G., and Pearson, P. M. An Evaluation of Ground Transport Requirements for Major Airports. Presented at Canadian Transportation Research Forum, May 1968.
9. Norhling, A. H. Future U.S. Transportation Needs. United Research, Inc., Cambridge, Mass., 1963.
10. Simpson, R. W. Future Short Haul Air Transportation in the Northeast Corridor. Seventh Annual Meeting of the Transportation Research Forum, 1966.
11. Koenig, H., et al. Analysis of Discrete Physical Systems. McGraw-Hill, 1961.
12. Seshu, S., and Reed, M. Linear Graphs and Electrical Networks. Addison-Wesley, 1961.
13. Roe, P. Networks and Systems. Addison-Wesley, 1966.
14. New York's Air Travellers. Port of New York Authority report to Eno Foundation, 1956.
15. Annual Report; 1967. Air Canada, Montreal, 1968.
16. Demand for Intercity Passenger Travel in the Washington-Boston Corridor. Systems Analysis and Research Corporation, Boston, Mass., 1965.
17. Airline Passenger O and D Statistics, 1964. Air Transport Board, Canada Department of Transport, Ottawa, 1965.
18. Canada Yearbook, 1967. Queen's Printer, Ottawa, 1967.
19. Population. In Census of Canada, 1961, Vol. 1, Part 2, Queen's Printer, Ottawa, 1962.

EFFECTS OF RESTRAINED HIGHWAY SPEEDS ON PROJECTIONS OF TRAVEL PATTERNS AND MODAL CHOICE

Anthony R. Tomazinis, University of Pennsylvania; and
Gerald J. Gaush, Delaware Valley Regional Planning Commission

This paper discusses the impact that use of restrained highway speeds may have in the simulation of future travel patterns. The investigation is made in 3 stages: at the time that choice of mode of travel is made, at the time of trip distribution, and at the time that automobile and transit trips are assigned to their respective networks. The comparison is made against the simulated travel pattern that is based on highway speeds previously determined as policy speeds. A review of the results indicates that a 15 percent increase in transit system accessibility produces approximately a 10 percent increase of transit trips in regions like the Philadelphia SMSA. This 10 percent increase in transit corresponds to approximately a 1.7 percent reduction in highway trips. The major transit diversion occurs in work trips and in central city trips. Most of the diversion (about 70 percent) took place within the central part of the region and along the corridor of the new rapid transit line from Philadelphia to Lindenwold.

●AS PART of the work program of the Delaware Valley Regional Planning Commission (DVRPC), an investigation was undertaken to ascertain the impacts of using the restrained highway speeds in projecting the 1985 travel patterns by mode of travel, by type of facility, and by subarea of travel within the region. The task involved the resimulation of the district-level travel projections for 1985 and their comparisons with earlier projections made by using the set of policy speeds that were adopted by the policy committee of the agency. In more specific terms, the resimulation required the resimulation of the modal split projections, distribution of the automobile and transit trips, and submodal split projections, and the reassignment of automobile and transit trips to the corresponding networks. The only phase of simulation process exempted was the trip generation; thus, total person trips generated for the region in 1985 remained unchanged at approximately 14.4 million.

Table 1 gives the differences between the policy speeds and the restrained speeds. The reference to policy speed for each facility of the highway network refers to the speed that was determined during the initial travel projections or the desirable speed on the basis of the type of the facility, its location, and a policy determination of what a desirable speed would be in each subarea and type of facility of the region in 1985. These speeds were developed by the staff and approved as policy inputs by the policy committee of the agency. In contrast, restrained highway speeds were obtained after the trip-assignment program of the DVRPC was run with the 1985 trip matrix. This trip-assignment program already includes a "capacity constraint" for the trip-assignment purposes (1). In some detail, the program accepts the trip interchanges, the link capacities, the policy speeds on each link, and a mathematical function that, at the end of each run, compares volumes to capacities by link and reduces correspond-

Sponsored by Committee on Transportation Forecasting.

TABLE 1

COMPARISON OF POLICY SPEEDS AND RESTRAINED SPEEDS
OF THE HIGHWAY NETWORK OF THE DVRPC REGION FOR 1985

Route Type	Area Type														
	CBD			Urban			Suburban			Rural			Open Rural		
	I[a]	II[b]	III[c]	I[a]	II[b]	III[c]	I[a]	II[b]	III[c]	I[a]	II[b]	III[c]	I[a]	II[b]	III[c]
Turnpikes	50	0	0	50	58	58	55	44	46	60	58	58	65	0	0
Non-Interstate freeways	37	39	39	44	38	39	50	44	45	55	55	55	60	0	0
Interstate freeways	37	36	35	44	37	40	50	43	45	55	53	54	60	0	0
Multilane, low type	19	12	12	23	15	16	34	23	24	39	34	34	45	0	0
Other	14	9	9	19	13	13	30	20	21	35	28	28	45	31	31
Multilane, high type															
Controlled access	26	20	21	31	24	27	41	34	35	47	41	43	50	0	0
Uncontrolled access	25	0	0	29	19	21	37	27	30	45	40	41	50	50	50

[a] Policy speeds (initial input).
[b] Restrained speeds (output of first round of assignment and input to restrained simulation of travel).
[c] Output speeds (at the final stage of assignments).

ingly the initial policy speeds on each link. The exact form of the function used is as follows:

$$Y = [M \cdot N \cdot (P + p)] + 1 - (M \cdot N \cdot P)$$

where

- Y = average daily restraint speed rate;
- M = daily duration of delay, $[2.5\ (v/c)^{-1}]^2$;
- N = peak-hour fraction, 2KD, obtained from look-up table based on route and area type;
- P = off-peak-hour policy speed rate, 60/off-peak speed in mph yields minutes per mile, obtained from look-up table based on route and area type;
- p = delay rate, $r(v/c)^5$;
- r = 2.5 for turnpikes, 3.0 for multilane high type, and 3.5 for other facilties;
- v = daily assigned volume; and
- c = daily capacity.

The program then repeats the process for the second, third, or fourth time, and so on. Because of numerous isolations of loadings that occur in many links after the third iteration, the impact of capacity constraints (of the form used by DVRPC) is only marginal. The final link loads are then estimated as the average of the loads that were produced by 2 sequential loadings (say, third and fourth iteration). The final link speeds are also the average speeds of the ones produced after the end of the 2 iterations that were used in estimating average loads.

PROJECTION OF TOTAL TRANSIT TRIPS

Modal-split projections involve essentially an estimation of the proportion (and subsequently the number) of future trips that will tend to make use of the public transportation system or the private automobile facilities. This estimation is accomplished usually by the use of a modal-split model. Such models are currently 1 of 2 types, i.e., either a trip-interchange model or a trip-end model (2). A trip-interchange model estimates the proportion of trips between 2 specific districts that would tend to make use of the transit system. This estimation takes place after the distribution of trips between districts. In contrast, a trip-end modal-split model estimates the proportion of total person trips generated in a district that would tend to make use of the transit system as a result of all the highway and transit connections of the district of trip origin. This estimation takes place before the distribution of trips between districts (3). This is the type of model used by DVRPC for reasons explained in other publications (4).

Table 2 gives the form that the DVRPC modal-split model took, together with the variables used and the statistical tests performed in its derivation. Table 2 gives some measures of quality of the model itself such as the correlation coefficient, the standard error, and the F-statistic. The correlation coefficients of the equations varied from 0.807 for non-work—home (NW-H) trips to 0.914 for work-home (W-H) trips from the 3 CBD's of the region. The standard error of estimate of the equations varied from 4.3 percent to 21 percent of the mean values of non-home—non-home (NH-NH) and non-work—home (NW-H) trips respectively. The F-statistic is also significant in all cases.

The statistical test of the significance of each of these variables indicated that not all variables are equally important nor is an important variable in 1 case necessarily important in another case. Nevertheless, the transit accessibility variable was found to be significant in several important cases. In the context of this paper, transit accessibility is defined as a measure of the relative ease of getting from 1 subarea to the entire region by using the transit system. This is also the variable that bears the impact of any major changes in speeds or additions of new major facilities in the highway or transit network.

TABLE 2

GENERATION OF TRANSIT TRIP ORIGINS IN EACH DISTRICT BY USING TRANSIT AND HIGHWAY SYSTEM VARIABLES

Trip Purpose	Area	Equation [a]	Correlation Coefficient R	Standard Error S	Statistic F	1960 Simulation Error	Standard Deviation Y	Mean $\bar{Y}$
Home-work	Cordon	$Y_{385} = 0.1626 - 0.4107 \log X_2 + 0.0709 \log X_{55} + 0.0702 \log X_{43} + 0.0001 X_{54}$	0.902	0.08	149.20	-0.82	0.18	0.26
Home—non-work	Cordon	$\log Y_{387} = -2.1544 - 0.7255 \log X_2 + 0.3314 \log X_{55} + 0.2142 \log X_{56} + 0.3058 \log X_{54}$	0.839	0.20	76.23	-7.19	0.35	0.08
Work-home	CBD[b]	$\log Y_{397} = -1.6234 + 0.1553 \log X_8 + 2.0913 X_{19}$	0.914	0.13	28.01	-0.81	0.21	0.43
	Cordon	$\log Y_{397} = -2.0513 + 0.2285 \log X_8 + 0.5700 X_{19} + 0.3104 \log X_{54}$	0.864	0.16	109.47	+4.17	0.25	0.19
Non-work—home	CBD[b]	$\log Y_{405} = -2.0556 + 0.6244 \log X_8 + 0.6143 X_{24}$	0.909	0.17	26.18	+8.11	0.25	0.32
	Cordon	$\log Y_{405} = -2.3800 + 0.4332 \log X_8 + 0.0371 X_{24} + 0.3198 \log X_{54}$	0.807	0.21	68.25	+3.50	0.35	0.07
Non-home—non-home	Cordon	$Y_{342} = -0.0257 + 0.0007 X_8 + 0.1971 X_{67}$	0.894	0.04	200.17	+4.82	0.09	0.07

[a] X_2 = Total cars per total occupied dwelling unit; X_8 = Jobs per nonresidential acre; X_{19} = Accessibility ratio for W-H trips; X_{24} = Accessibility ratio for NW-H trips; X_{43} = Accessibility ratio for H-W trips; X_{54} = Average transit vehicles per day; X_{55} = Permanent occupied dwelling units per residential acre; X_{56} = Accessibility ratio for H-NW trips; X_{67} = Accessibility ratio for NH-NH trips.

[b] Philadelphia, Camden, and Trenton.

The significance of the accessibility variable was tested statistically in terms of the t-test and the beta-test of the coefficients of the equations. The results have shown that its significance is different for each type of trip. For H-W and H-NW trips, the absolute contribution of transit accessibility is rather small. Clearly the speed of the transit system has a very small association with the number of H-W and H-NW trips made by transit. For W-H trips from the 3 CBD's, the situation is reversed. It appears that people going home from work and from a CBD area place special emphasis on a fast transit system. This is clearly shown by the beta-values and the t-test of the coefficients. For W-H trips from non-CBD origins, however, the significance of the transit accessibility variable is much smaller than either the job density or the frequency of transit service variable. For NW-H trips from the 3 CBD's of the region or for the same trips originating from elsewhere in the region, the transit accessibility variable is also found to have little association with the rate of transit trips. Finally, for NH-NH trips, the transit accessibility (or speed and cost of the transit system) also emerges as a relatively important factor as indicated by both pertinent statistical yardsticks.

Before the accessibility variable was used in the modal-split model, a comparison was made of the 2 sets of accessibility measures (i.e., with policy speeds versus restrained speeds). The results of these comparisons indicate that (a) the largest changes in the accessibility measure occurred for districts within the central part of the region (primarily within Philadelphia, then in Camden, and then in Trenton); (b) the H-W trips register the greatest increases, whereas the NH-NH trips register the smallest increases; and (c) an increase in the transit accessibility has been registered for almost all districts in the region for all 5 trip purposes. No district has experienced a decrease of transit accessibility, and the change has shifted an average line of accessibility to the right of a 45-deg line between restrained and unrestrained accessibility. This shift corresponds, in turn, to approximately 25 percent increase of accessibility for low-accessibility districts to almost 12 percent increase of accessibility for high-accessibility districts.

The application of these changes to the modal-split model of DVRPC produced the new regional projections of transit and highway trips given in Table 3. It becomes apparent that the city of Philadelphia (sectors 1 through 6, Fig. 1) registers the largest shift of trips from highway to transit. Almost 72 percent of the total new diversion to transit occurs in the city. The Philadelphia CBD (sector 1) experiences a greater shift than any other area of the city, accounting for almost 30 percent of the total shift in the region. This appears to be a reasonable result if one considers the fact that the impact of restrained speeds would be felt more strongly on those highway facilities in the CBD than in any other part of the region. At the same time, the CBD is the hub of the region's transit network (rapid transit included), and, as such, the reduction of highway speeds would produce the most striking contrast with the transit speeds.

Also noteworthy is the increase in transit trips in sector 11. This sector contains the modern Lindenwold Rapid Transit Line, which accounts for the fact that this sector absorbs almost 10 percent of the regional total increase in transit trips. The high speeds and quality of service of this line make it clearly an attractive alternative when compared to slowly moving highway facilities resulting from restrained speeds.

The results indicated that the work-trip purpose (H-W and W-H) registers the largest decrease in highway trips—61 percent of the total diversion to transit. If one recalls the changes of accessibility measures presented earlier, these results become both expected and reasonable.

On a regional basis, the total decrease in highway trip origins was less than 2 percent. (Highway trips with either external origins or destinations and through trips were assumed to be unaffected by restrained speeds. This was done because the highway speeds and the transit service within the cordon are expected to have minimal impact on through trips or on trips that only partly use the regional network, i.e., external-internal trips.) The effect of restrained speeds on the regional highway network is, therefore, not a dramatic one. Specific corridors that may register greater changes are explored in the next section of this paper, but it is important to notice that total regional changes of highway trips were very small.

TABLE 3

RESTRAINED AND UNRESTRAINED TRIP PROJECTIONS BY SECTOR FOR 1985

Sector Number	Intermediate Highway[a]				Basic Transit[b]			
	Policy Speeds	Restrained Speeds	Difference		Policy Speeds	Restrained Speeds	Difference	
			Trips	Percent			Trips	Percent
1	352,421	313,467	38,954	-11.1	323,879	376,060	52,181	+16.1
2	466,674	454,130	12,544	-2.7	241,622	257,896	16,274	+6.7
3	278,135	272,202	5,933	-2.1	108,134	116,562	8,428	+7.8
4	563,561	551,795	11,802	-2.1	239,693	256,223	16,530	+6.9
5	589,933	575,232	14,701	-2.5	228,302	247,324	19,022	+8.3
6	607,037	597,547	9,490	-1.6	153,359	165,632	12,273	+8.0
7	600,592	598,461	2,131	-0.4	27,642	30,392	2,750	+9.9
8	842,380	835,258	7,122	-0.8	77,774	87,185	9,411	+12.1
9	1,401,694	1,390,594	11,100	-0.8	169,581	184,509	14,928	+8.8
10[c]	444,232	441,221	3,100	-0.7	47,801	52,003	4,202	+8.8
11[d]	1,190,345	1,177,732	12,613	-1.06	148,977	165,772	16,795	+11.3
12	488,290	485,727	2,563	-0.5	36,007	39,275	3,268	+9.1
Total	7,825,294	7,693,330	131,964	-1.7	1,802,771	1,978,833	176,062	+9.8

[a] Vehicle trips.
[b] Person trips.
[c] Includes Sector 16, Trenton CBD.
[d] Includes Sector 15, Camden CBD.

The percentage increase in transit trips on a regional basis is, of course, more substantial, reaching almost 10 percent of the original projections for this combination of networks. Again, what this regional change implies in specific corridors will be presented in a following section of this paper.

TRIP DISTRIBUTION

Trips from the restrained speeds modal split were distributed by the gravity model as used by DVRPC. Inputs such as friction factors, K-factors, bridge penalties, parking costs, and intradistrict costs were identical to those used in the original 1985 district trip distribution using policy speeds. For transit trips, the only difference was in the number of trips distributed (increase by 176,000 trips).

For highway trips, the difference in distribution was both in the number of trips distributed (decreased by 176,000 person trips or 132,000 vehicle trips) and in the new minimum cost paths that were produced by using the restrained speeds.

Table 4 gives a summary of the results of the 2 trip distributions for 1985 highway and transit trips. The total highway trips, as already mentioned, decrease by 132,000 vehicle trips (equivalent to 176,000 person trips after the application of the car-occupancy factors by trip purpose), but also the number of intradistrict highway trips increase by about 376,000 vehicle trips. Intradistrict trips stay rather constant for transit. In terms of mean trip cost, the highway trips generally increase in travel cost (most noticeably for the work purposes), whereas the transit trips stay rather stable in cost.

The gravity distribution model is, of course, synthetic in nature, determining choices of destinations in a relative manner based on relative travel costs from a given origin point. Thus, by increasing travel cost (using restrained speeds), we increase the friction of space to all destinations outside the district of origin and, therefore, we set in a more favorable position the destinations located within the district of trip origin itself. This situation results in the production of many more intradistrict highway trips (17 percent more), and this, in turn, has significant impacts on all indexes of trip distribution (such as mean trip cost and total travel cost). Because no change in cost paths is introduced in the transit network, no reason exists for changing the proportion of interdistrict transit trips. In fact, the distribution registers only 6.0 percent increase of intradistrict transit trips (versus 10 percent increase of total transit trips), which tends to explain the slight increase of mean trip cost of transit trips registered.

The structure of the gravity model and the nature of the differences introduced in the distribution help to explain the relatively large increase in mean trip cost of highway trips registered, especially for work trips. Trips have to cross congested parts of the networks with unit travel cost being much higher than the cost based on policy speeds. Because work trips have more or less restricted destinations, it should be expected that their trip cost would increase. This is exactly what happened, especially for H-W and W-H trips. For NH-NH trips, this situation did not materialize at all, and the mean trip cost actually decreased slightly.

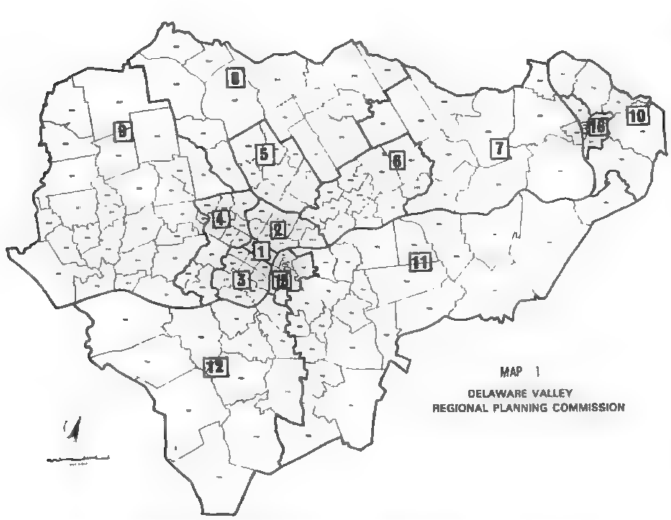


Figure 1. Traffic simulation sectors within cordon area grouped by data collection district (DVRPC).

TABLE 4

SUMMARY OF RESTRAINED AND UNRESTRAINED INTERNAL TRIP DISTRIBUTION

Trip Purpose	Network[a]	Speed	Total Trips Distributed	Intradistrict Trips	Mean Trip Cost (cents)	River Crossings
H-W	Highway	Policy	915,503	118,076	68.48	97,883
		Restrained	875,537	137,946	72.84	87,297
	Transit	Policy	444,318	15,456	91.07	48,106
		Restrained	494,133	16,450	91.40	54,672
H-NW	Highway	Policy	2,026,300	661,063	41.85	59,798
		Restrained	2,012,722	776,145	42.24	56,941
	Transit	Policy	363,112	30,307	80.14	25,904
		Restrained	386,380	31,266	79.59	26,613
W-H	Highway	Policy	914,755	130,477	68.28	96,183
		Restrained	873,197	151,939	73.23	86,172
	Transit	Policy	443,483	15,301	91.37	48,952
		Restrained	494,097	16,306	91.60	55,539
NW-H	Highway	Policy	2,049,611	680,679	41.66	56,870
		Restrained	2,035,523	794,063	41.74	50,519
	Transit	Policy	363,983	30,433	78.99	24,331
		Restrained	387,899	31,353	78.82	25,648
NH-NH	Highway	Policy	1,919,125	588,460	37.94	47,142
		Restrained	1,896,351	694,456	37.28	38,995
	Transit	Policy	187,875	11,941	78.07	18,610
		Restrained	216,324	14,449	78.89	21,318
Total	Highway	Policy	7,825,294	2,178,755		375,876
		Restrained	7,693,330	2,554,549		319,924
	Transit	Policy	1,802,771	103,438		165,903
		Restrained	1,978,833	109,824		183,790

[a]Vehicle trips on highway network, and person trips on transit network.

Table 4 also gives a significant change in the Delaware River crossings. The highway crossings show a decrease of about 38,000 vehicle trips, which is a 10 percent loss resulting from restrained speeds. Converted to person trips, this becomes about 52,000 trips, which is considerably higher than the approximately 18,000 gain in transit river crossings. The increase of the friction of space in crossing the river (using congested highways) has again forced the gravity model to register a reduction of trips selecting this direction and to register a diversion of vehicle trips to destinations on the same side of the Delaware River. These destinations become more attractive substitutes because one does not have to use so many congested highways to reach them.

Mean trip costs for all purposes of transit have remained fairly stable under application of restrained speeds. This result occurs despite the 10 percent increase in total trips. Indications are, therefore, that the 1985 transit network will be more efficiently used when trips are diverted from the highway network and that it will handle the increased volume without any significant increase in average trip cost.

Delaware River crossings by transit have increased about 11 percent (18,000 person trips corresponding to approximately 14,000 automobile trips) as a result of applying restrained speeds to the highway network. Mainly, this is attributable to the

TABLE 5

RESTRAINED AND UNRESTRAINED TRIP PROJECTIONS TO AND FROM THE PHILADELPHIA CBD

Sector	Highway Vehicle Trips			Transit Person Trips		
	Policy Speed	Restrained Speed	Percentage Change	Policy Speed	Restrained Speed	Percentage Change
From Sector and to CBD						
1	18,199	22,621	+24.3	10,563	11,211	+6.1
2	38,112	35,646	-6.5	33,508	38,833	+15.9
3	47,357	52,132	+10.1	26,475	31,088	+17.4
4	62,599	59,024	-5.7	50,882	59,177	+16.3
5	25,175	19,314	-23.3	51,143	59,703	+16.7
6	21,972	19,442	-11.5	36,339	42,711	+17.5
7	11,365	10,008	-11.9	8,096	9,474	+17.0
8	25,549	19,902	-22.1	21,518	25,350	+17.8
9	71,558	55,057	-23.1	46,099	52,934	+14.8
10[a]	248	306	+23.4	609	625	+2.6
11[b]	24,255	17,872	-26.3	32,211	36,920	+14.6
12	10,028	7,631	-23.9	7,282	8,608	+18.2
Total	356,417	318,955	-10.5	324,725	376,634	+16.0
Difference	-37,300			+52,000		
From CBD and to Sector						
1	18,199	22,261	+24.3	10,563	11,211	+6.1
2	32,949	30,268	-8.1	37,955	43,894	+15.6
3	42,555	46,315	+8.8	27,382	32,124	+17.3
4	65,143	61,263	-6.0	51,624	60,031	+16.3
5	31,308	24,903	-20.5	50,021	58,510	+17.0
6	22,074	19,440	-11.9	34,167	40,352	+18.1
7	10,324	8,799	-14.8	7,323	8,603	+17.5
8	24,664	19,056	-22.7	19,913	23,597	+18.5
9	70,876	54,002	-23.8	47,811	54,838	+14.7
10[a]	169	208	+23.1	422	467	+10.7
11[b]	23,997	18,757	-21.8	29,652	34,119	+15.1
12	10,157	7,833	-22.9	7,043	8,296	+17.8
Total	352,415	313,465	-11.1	323,876	376,042	+16.1
Difference	-39,000			+52,000		

[a] Includes Sector 16, Trenton CBD.
[b] Includes Sector 15, Camden CBD.

effectiveness of the Lindenwold Line in draining trips from the highways in its corridor and carrying them across the Benjamin Franklin Bridge into Philadelphia. As was pointed out before, a rapid transit facility such as this cannot be anything but favorably affected when restrained speeds are applied to the highway network. Finally, the increase in river crossings is in line with the increase in total transit trips regionally and seems reasonable on all counts.

Another significant aspect of the distribution, worthy of discussion at this juncture, is the nature of travel patterns to and from the Philadelphia CBD. These patterns are given in Table 5. Overall, there is a decrease in highway vehicle trips to and from

TABLE 6

SUMMARY OF RESTRAINED AND UNRESTRAINED TRIP ASSIGNMENT FOR 1985

Item	Policy Speeds	Restrained Speeds	Difference	
			Amount	Percent
Total trips assigned	8,900,000	8,400,000	131,945	-5.6
Vehicle-miles, daily	55,481,349	50,575,920	4,905,429	-8.8
Vehicle-miles, peak hour	6,532,558	5,973,910	558,648	-8.6
Vehicle-hours, daily	2,229,263	1,924,627	304,636	-13.7
Miles of route	2,590	2,590	—	
Average speed, mph	24.9	26.3	1.4	+5.6
Vehicle time cost, $/day	3,343,894	2,886,939	456,955	-13.7
Vehicle operating cost, $/day	1,845,880	1,666,405	179,475	-9.7
Accident cost, $/day	696,741	580,546	116,195	-16.7
Toll cost, $/day	186,366	176,944	9,422	-50.6
Total cost, $/day	6,072,882	5,310,834	762,048	-12.5
Average trip cost, cents	54.7	48.5	6.2	-11.3
Cost per vehicle-mile, cents per mile	10.9	10.5	0.4	-3.7
Average trip length, mile	5.02	4.62	0.4	+8.0

Note: These are regional summaries and do not include approach links.

the CBD of approximately 38,000 trips or about 11 percent. On the other hand, there has been an increase in CBD-oriented transit person trips of approximately 52,000 trips or about 16 percent of the original figures resulting from using policy speeds. In addition, almost 60 percent of the total increase in transit trips to and from the CBD occurs with connection to sectors of the city of Philadelphia (sectors 1 through 6). If one considers that most of the impact of the restrained speeds occurs on the network within the city of Philadelphia, this finding should not be surprising.

TRAFFIC ASSIGNMENT

Highway Trip Assignment

The traffic assignment phase is the final step in the simulation process. Completion of the assignment program provides more detailed data such as link volumes, bridge volumes, travel costs, facility speeds, and other measures of network efficiency that permit a more rational evaluation of the need and role of each facility (5, 6).

Summary results of the district highway traffic assignments are given in Table 6. The restrained assignment indicates a reduction in all summary categories from the ones produced with the policy speed assignment. It should be noted that the total trips assigned represent the actual trips assigned on the network, that is, after the intradistrict trips obtained from the trip-distribution process were deducted from total trips to determine the number of trips actually to be assigned. The total reduction amounted to approximately 500,000 vehicle trips (i.e., 132,000 vehicle trips reduced due to modal-split shifts and 376,000 vehicle trips that became intradistrict trips in addition to the intradistrict trips produced on the basis of policy speeds.)

The regional cost summaries show that the restrained speed assignment experiences a significant (almost 14 percent) reduction in cost from the policy speed levels. This is an obvious result of loading 5 percent fewer trips on the highway network. Also, the average trip length (in dollars and miles) of the assigned highway trips has been reduced. Clearly, the majority of diverted trips are among the longest ones in the region. The combination of these 2 factors produces fewer total vehicle-miles of travel,

TABLE 7

RESTRAINED AND UNRESTRAINED VEHICLE-MILES BY ROUTE FOR 1985

Route Type[a]	Policy Speed Assignment	Restrained Speed Assignment	Change (percent)
Turnpikes	3,885,894	3,717,887	-4.3
Freeways	22,782,662	20,909,344	-8.2
High type of arterials	6,120,630	5,791,718	-5.4
Low type of arterials	22,692,163	20,156,971	-11.2
Total	55,481,349	50,575,920	-8.8

[a]Excludes approach links.

slightly higher average speeds for the entire network, and, in turn, lower costs per vehicle-mile and lower total operating costs to the users.

Restrained speed assignments have varying degrees of impact on different route types. Comparative analysis of this effect is given in Table 7. Turnpikes register the least effect, experiencing only a 4.3 percent decrease in vehicle-miles of travel. A substantial portion of travel on turnpikes is long-distance travel and not very susceptible to any localized diversion to transit; therefore, the foregoing result appears quite reasonable. The opposite end of the spectrum, however, is represented by the low type of arterials. This route type undergoes an 11.2 percent reduction of vehicle-miles in restrained speed assignment. The low type of arterials serves the shorter trips and suffers the most from congestion and speed reduction and, as a consequence, will reflect the greatest diversion to transit.

Another interesting facet of the assignment is the impact of restrained speeds on the projected volumes on the Delaware River bridges. In total, there is a reduction of almost 38,000. The interesting point to note is, On which bridges do the greatest reductions occur? The largest decreases in volume occur on the Benjamin Franklin, Walt Whitman, and Philadelphia-Pennsauken bridges. These facilities are within the area of influence of the Lindenwold Line and are served by relatively congested facilities. As a result, they experience substantial diversion to the high-speed line and to other bridges. Five other bridges actually are projected to have slight increases in volume, which is due to their geographic locations. The locations are such that they have no proximity to a practical transit alternative, are in uncongested areas, and are appropriate for receiving diversions from other congested bridges.

Certain highway facilities in the region are worth examining in a little more detail. The Crosstown Expressway restrained speed assigned volumes consistently show a decrease from policy speed levels on all links. Also, the Vine Street Expressway generally shows a decrease, although some links actually increase in volume. Both of these facilities are affected by the Market Street Subway, which runs in a parallel direction, and by the Locust Street Line and the many bus routes serving the city. The observed volumes on the Schuylkill Expressway show an irregular variation by link. Whenever there is an alternative link, the expressway links decrease in volume because this expressway is exceedingly congested throughout the day. The Delaware Expressway shows a consistent and significant decrease in volume reflecting the decrease of highway trips generated in its vicinity and the congested nature of the facility's operation.

At the conclusion of the highway trip assignment, the final network speeds were averaged and reported. A comparison is included in data given in Table 1 between these final speeds and the ones used as inputs for the modal split, trip distribution, and highway trip assignment (the restrained highway speeds). As one can notice in many cases, the final restrained speed was higher than the input restrained speed. In fact, from the 35 cases presented, 18 cases remain the same, 16 cases increase, and

TABLE 8

1985 TRANSIT TRIP ORIGINS BY SUBMODE BY SECTOR

Sector	Railroad Trips			Subelevated Trips			Surface Bus Trips		
	Policy Speed	Restrained Speed	Percent Increase	Policy Speed	Restrained Speed	Percent Increase	Policy Speed	Restrained Speed	Percent Increase
1	56,300	65,300	15.9	142,200	166,500	16.3	125,400	142,200	14.9
2	3,900	4,200	7.6	85,700	91,700	7.1	152,000	162,000	6.6
3	500	500	0.0	41,200	44,100	7.0	66,400	71,900	8.2
4	7,000	7,500	7.1	85,100	91,400	7.4	147,600	157,300	6.6
5	12,300	14,200	15.4	91,800	100,900	9.8	124,300	132,300	6.5
6	2,500	2,900	16.0	80,500	89,000	10.5	70,300	70,700	4.8
Total Philadelphia	82,400	94,500	14.7	526,500	538,800	11.0	686,000	741,400	8.2
7	6,600	7,600	15.2	7,200	8,000	11.1	13,800	14,800	7.3
8	20,800	24,300	16.8	13,100	14,600	11.3	43,800	48,300	10.0
9	30,100	33,900	12.5	43,500	47,500	9.2	96,000	103,100	7.3
Pennsylvania suburbs	57,600	65,800	14.3	63,800	70,100	9.7	153,600	166,200	8.3
Total Pennsylvania	140,000	160,300	14.5	590,300	653,900	10.8	839,700	907,600	8.2
10	200	200	0.0	900	900	0.0	15,500	16,500	6.4
11	900	1,000	11.1	24,900	28,000	12.5	61,800	68,900	11.4
12	200	200	0.0	6,800	7,300	7.3	29,000	31,800	9.7
15	400	400	0.0	13,200	14,600	10.6	47,800	52,900	10.7
16	400	400	0.0	1,200	1,200	0.0	29,600	32,700	10.4
New Jersey suburbs	2,100	2,200	4.7	46,900	52,000	10.7	183,700	202,900	10.0
Total region	142,200	162,500	14.3	637,200	705,900	10.8	1,023,300	1,110,400	8.4

Note: Figures may not add because of rounding.

only 1 case decreases (for Interstate freeways in the CBD area). Recalling the fact that the input speeds are already quite restrained, because of the inclusion of the volume-capacity ratio delay function in the assignment phase of the initial traffic simulation, one can understand why the final highway speeds increase when the total impact of restrained speeds is incorporated in the modal split, trip distribution, and trip assignment.

Transit Trip Assignment

The task at this part of the project involves essentially the assignment of the new total transit trips, projected for 1985, on the 3 transit subsystems of the region. In reality, the transit trip assignment in the DVRPC package requires that the analyst first go through the submodal-split process. The purpose of this step of the work program is to determine the number of trips to be assigned to the 3 public transportation subsystems, that is, the railroads, the subway-elevated, and all the surface bus lines. [The submodal split is performed in 3 steps. First, the number of railroad trips are computed by applying diversion curves derived from 1960 data on railroad passengers to the total public transportation trips. The second step uses diversion curves derived from 1960 subway-elevated ridership, applies them to transit passengers (total public transportation less railroad), and obtains subway-elevated trips. The remaining trips represent all surface riding (7).]

The results of the transit trip assignment reveal again significant differences from the results obtained with the previous estimate of transit trips. These differences are given in Table 8 on a sector basis for both policy and restrained speed runs. Again, the importance of the differences in the city of Philadelphia is obvious, as are those for sector 11, which includes the city of Camden and the Lindenwold Line. With regard to submodal comparisons, it becomes clear that the railroad system receives proportionally more trips from the newly diverted trips to the transit system. For the region, the railroad system receives 14.3 percent more trips versus 10.8 percent more trips for the subway system and only 8.4 percent more trips for the surface bus lines. Furthermore, sectors with good railroad connections are clearly the ones with the greatest proportion of diverted trips to the railroad system.

Ridership projections for each individual transit facility, produced by both policy and restrained highway speeds, were posted on each line by link and station. Maps were then prepared, and results were tabulated for further economic analysis of each line.

SUMMARY AND CONCLUSIONS

The overall results have shown that for volumes on the highway network the effects of restrained speeds are very small. The change in highway trips was not enough to indicate that the numbers of automobile drivers who would be discouraged by congestion and decide to use transit are sufficient to either reduce a single planned expressway facility or add a new transit line.

Transit projections registered a 10 percent increase in the regional ridership, but the significance of this increase is rather small when one realizes that it is based on a relatively small number of trips and is going to be served by a transit line that can hardly be considered overloaded. The size of such an increase and its areal distribution reduces any chance that it might conclude that a new transit system is needed at a specific location. The corollary statement can also be made; i.e., changes in the transportation system characteristics must be quite substantial before any significant effect (above 10 percent) will be noticed in the number of transit trips. It should be also pointed out that this paper reports only on the effects of a relative downgrading of highway network quality. The transit network, as proposed, was left unchanged. Perhaps stronger insights could be gained from testing the effects of combined changes, e.g., restrained highway speeds and greater frequency of transit service. Because there is currently a renewed interest in transit, and particularly in the newer, more exotic forms, it would be interesting to incorporate system changes of this type into the network and test the effects on modal preference. Present models, however, are not

structured to cope with this type of problem. Nevertheless, it is the type of tests that should be the focus of future efforts, if forecasting future transit use is to be a useful tool to the transportation planner in revitalizing transit and alleviating highway congestion.

The new distribution of automobile driver and transit trips reveals important variations. First the intradistrict automobile driver trips register a marked increase (17 percent). This increase of intradistrict trips occurs primarily in NW-H and H-NW trips (61 percent of the increase) and in NH-NH trips (28 percent of the increase). A second important change was registered in the general increase of trip cost found for all automobile driver trips to or from home, and particularly for H-W and W-H automobile driver trips. For all transit trips and for NH-NH automobile trips (the shortest automobile trips), no significant difference in trip length, measured in total travel cost, was found. A third important change was found for trips crossing the Delaware River. In general, a reduction of 38,000 automobile driver trips (corresponding to 52,000 automobile person trips) and an increase of 18,000 transit trips (11 percent of the total diversion to transit) were observed. The reduction of automobile driver trips crossing the Delaware occurs primarily in H-W trips (10,000), W-H (10,000), and NH-NH trips (9,000). The increase in transit trips occurs primarily in the same types of trips (i.e., H-W has an increase of 6,000 trips, W-H has an increase of 6,000 trips, and NH-NH has an increase of 3,000 trips). In total, 34,000 automobile person trips change destination and do not cross the river.

The restrained speeds have also produced an important change in the travel pattern to and from the Philadelphia CBD. In general, a reduction of 52,000 automobile person trips was registered (corresponding to 38,000 vehicle trips), and an equivalent increase of transit trips was reported. The transit trips increase corresponds to 16 percent of previous projections and comes primarily from districts located increasingly distant from the Philadelphia CBD. Also, the sector that includes the new Lindenwold Rapid Transit Line registers the largest percentage decrease of automobile trips to the Philadelphia CBD. The increase of transit trips to the Philadelphia CBD from this sector is among the largest ones, reflecting the fact that the previous projections from this sector already include the impact of the new transit line.

The assignment of vehicle trips on the restrained-speed network reveals that the reduction of 500,000 vehicle trips (132,000 trips diverted to transit plus 376,000 trips becoming intradistrict trips and therefore not assignable) produces a reduction of 4.9 million vehicle-miles and 304,000 vehicle-hours. Most of this reduction occurs on low types of arterials (-11.2 percent), which in turn causes an increase in the average travel speed from 24.9 to 26.3 mph and a reduction of average trip cost from 54.7 to 48.5 cents, average trip length from 5.02 miles to 4.62 miles, and cost per vehicle-mile from 10.9 to 10.5 cents per mile. It is clear that the trips diverted from the highway network are primarily those previously forced to use an indirect path on local arterials. In terms of river crossings, assignment of vehicle trips reveals that the reduction of 38,000 vehicle trips registered in the trip-distribution phase of the comparison is affecting the various bridges differently. Diversions have also been observed from 1 bridge to another with less congested facilities.

In terms of the effect of restrained speeds on the expressway networks, the most uniform impact is observed on the Crosstown Expressway where the assigned traffic load is reduced substantially (30 percent on 1 link). Similar reductions occur on other congested parts of the major networks—Vine Street and Schuylkill Expressways.

This restrained-speed analysis also has implications insofar as regional development patterns are concerned. Congested, slowly moving highways cause diversion to transit and thus induce a somewhat more compact development of the central part of the region. Also, for suburban residents who find that a shift to transit for travel to the Philadelphia CBD is not a desirable or viable alternative, the congested highway speeds constitute a good inducement to avoid long trips and to seek travel destinations close by or in uncongested areas (i.e., mostly suburban areas). To this extent a congested highway system in the central part of the region becomes a strong influence for greater suburbanization and more self-supporting suburbs.

ACKNOWLEDGMENT

The authors wish to acknowledge the contributions made by several members of the DVRPC Transportation Division staff in the preparation of this paper. Special thanks are given to Richard E. Hubbell and William Weigand for their constructive criticisms and editing. Calculations and indexes used by this paper were completed under the supervision of Frank McHugh, Arnold Pinsley, and Jose Berguido. The responsibility for all statements and conclusions included herewith rests with the authors alone.

REFERENCES

1. Wickstrom, G. V. Basic Relationship Between Daily Work Travel and Peak-Hour Traffic. Penn-Jersey Transportation Study, Tech. Memo. 13-9, 1964; Traffic Engineering, Vol. 34, No. 5, 1964.
2. Fertal, M. J., et al. Modal Split. U.S. Department of Commerce, Bureau of Public Roads, Dec. 1966.
3. Transit Usage Forecasting Techniques: A Review and New Directions. Consad Research Corp., prepared for the U.S. Department of Housing and Urban Development, Pittsburgh, Penn., April 1968.
4. Tomazinis, A. R. Modal Split Model in the Penn-Jersey Transportation Study Area. Highway Research Record 165, 1967, pp. 41-75.
5. 1985 Alternative Test No. 4, Estimated Effect of Restrained Highway Speeds on Transit Ridership Projections. Prepared for Technical Advisory Committee on Highway Plans and Transit Plans of the Delaware Valley Regional Planning Commission, Philadelphia, Penn., May 1970.
6. Zakaria, T., Sr., and Falcocchio, J., Sr. The Traffic Assignment Process of the Delaware Valley Regional Planning Commission. Highway Research Record 250, 1968, pp. 35-53.
7. Pinsley, A. G. The Sub-Modal Split Model. Prepared for Planning Coordinating Committee of the Delaware Valley Regional Planning Commission, Philadelphia, Penn., Sept. 27, 1967.

CHOICE AND CAPTIVE MODAL-SPLIT MODELS

Michael G. Ferreri, Simpson and Curtin, Transportation Engineers; and
Walter Cherwony, Wilbur Smith and Associates

This paper presents the findings of an investigation conducted to determine the desirability and feasibility of developing 2 sets of modal-split models—direct generation for captive transit riders and trip interchange for choice transit riders. The different characteristics of the dual transit trip-making population are defined by a set of hypothetical demand curves of transit services for both captive and choice riders. Whereas captive riders are inelastic with changes in the relative attractiveness of the transit and highway systems, choice riders are sensitive to changes in either highway or transit facilities. The standard method of treating these groups as one usually produces equations that are insensitive to system variables because the characteristics of the captives tend to dominate the models. Information gathered for the Miami, Florida, urban area was used to identify and quantify those factors affecting transit use into 2 sets of models.

•A CONSIDERABLE AMOUNT of effort has been expended during the past decade by transportation planners to solve the recurring question of what the division is likely to be between highway and transit use for a significantly different future transportation system. Most reviews of past research catalog all modal-split models into predistribution, post-distribution, and direct generation models. Stratification of transit rider data is usually by trip purpose. A common ailment of most of these efforts—especially in cities without existing rapid transit—is the lack of proof that relationships devised from existing conditions are meaningful when extrapolated into the future with major changes in the transit system.

In a recently completed examination of the state of the modal-split art, the CONSAD Research Corporation (1) concluded that "the present situation . . . finds none of the models producing forecasts which can be reliably used for decision-making purposes where major system changes are contemplated." A continuing problem is to develop models that are statistically sound but sufficiently sensitive to changes in the transportation system to reflect the effect of new transit modes.

Examination of transit patrons always reveals 1 consistent dichotomy regardless of the urban area: They subdivide into captive riders, who have no alternate means of transportation, and choice riders, who have an automobile available and could use it if they desired. In practically every city, captive riders outnumber choice riders by ratios ranging anywhere from 3:1 to 9:1. Captive riders as a group are inelastic with changes in the attractiveness of the transit system. If they are to make the trip at all, it must be made by transit. For this reason, attempts to develop models with system variables are futile insofar as this major group is concerned. The fact that they outnumber choice riders by such wide margins overrides any sensitivity that choice riders may have to changes in the transit system when the 2 groups are lumped together for multiple regression analysis.

On the other hand, choice riders in any urban area are using the transit system because it provides some distinct advantage (by their evaluation) over traveling by automobile. If anything is to be learned about responses of travelers to alterations in the

Sponsored by Committee on Transportation Forecasting.

relative attractiveness of the transit and highway networks, it must be learned from this group of transit users.

For these reasons and the fact that prior attempts at the development of typical trip interchange modal-split models in Metropolitan Dade County (Miami, Florida) produced results that were completely unresponsive to changes in the transportation system, it was decided to conduct the research reported here that examined the 2 basic groups of transit users in Miami—captive and choice—and that resulted in the development of separate modal-split techniques for each group.

Preliminary review of the characteristics of these groups and early regression analysis indicated not only that the groups should be divided but also that the model theory should be applied differently to each group. Tests were conducted that led to the development of direct generation models for captive riders and trip-interchange models for choice riders.

The rationale in selecting a direct-generation model for captive riders and a trip-interchange equation for choice riders can be explained by examination of a set of hypothetical demand curves (Fig. 1) for the 2 individual trip-maker populations. The graph relates the demand for transit service (expressed as the percentage of the trips made by transit) to a disutility index. This index quantifies the costs associated with each travel mode as measured by elements such as travel time, out-of-pocket expenses, and comfort and convenience factors. The index ranges from 0, which indicates almost ideally attractive transit service in relation to highways, to 1, which denotes equality in service, up to a maximum of 10, which represents an ideally attractive highway system (the opposite of the 0 condition).

For the captive trip-makers, the demand function is inelastic to the disutility index of the transportation system, and all trips are made by transit regardless of the value of the index. Of the 3 types of modal-split models (direct generation, trip end, and trip interchange), the direct generation procedure is the least sensitive to changes in the transportation system. Therefore, it was the most compatible with the captive transit demand curve. On the other hand, the choice trip-makers respond to changes in either highway or transit services as reflected in the disutility index. The elasticity of the transit demand function for choice riders can best be modeled by a trip-interchange type of formula because it is most responsive to specific system changes.

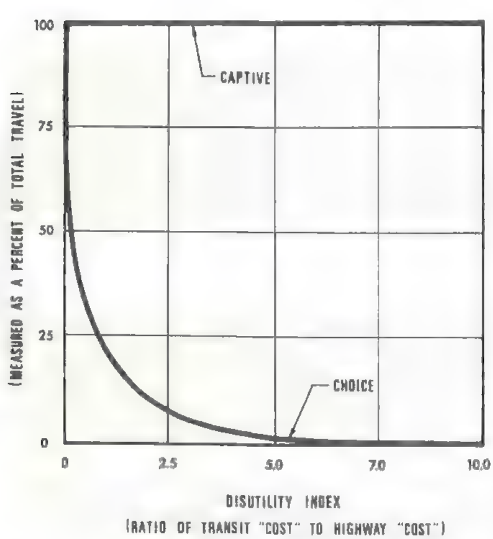


Figure 1. Hypothetical transit travel-demand curves for captive and choice riders.

CAPTIVE TRANSIT RIDER MODELS

As previously noted, the captive transit riders represent those members of the trip-maker population who have no automobile available for their trips and, therefore, have no choice in the selection of travel modes. If the trip is to be made, it must be made by transit. In view of this characteristic, the possibility of developing direct transit trip-generation models was examined. With this method, the number of captive transit trips produced and attracted in each zone is related to the demographic and economic characteristics of the individual zone. Six different trip categories were chosen to reflect the different characteristics of travel by trip purpose: home-based work, home-based shop, home-based social and recreational, home-based school, home-based miscellaneous, and non-home-based. These trip-purpose categories are the same as those developed by the Miami Urban Area Transportation Study (MUATS) for estimating

TABLE 1

INPUT VARIABLES FOR REGRESSION ANALYSIS—CAPTIVE TRANSIT TRIP-ESTIMATING EQUATIONS

Variable	Mnemonic	Variable	Mnemonic
Residential dwelling units	RESDU	Labor force	LABFOR
Hotel and motel units	HMUNIT	Eating, drinking and amusement recreation employment	ED + AR
Total dwelling units	TOTDU	General merchandise employment	GMEMP
Resident population	RESPOP	Commercial employment	COMEMP
Tourist population	TOUPOP	Total employment	TOTEMP
Total population	TOTPOP	Junior high school enrollment	ENJRHS
Automobiles owned by residents	RESAUT	Senior high school enrollment	ENSRHS
Automobiles owned or rented by tourists	TOUAUT	Other enrollment	ENOTH
Total automobiles owned or rented	TOTAUT	College enrollment	ENCOLL
		Total home-based attractions	HBATTR

total person trips. In a sense, the captive transit estimating equations can be thought of as minimodels that forecast only a segment of the total person trips.

Because certain portions of the 550-zone study area are not currently served by transit, only 452 zones or cases were used in model calibration. Generally, the excluded zones are located at the periphery of the study area and are largely undeveloped. A total of 19 independent variables (Table 1) were correlated with each of the 12 dependent travel variables (trip productions and attractions for each of the 6 trip purpose categories) by using the BMD02R statistical program. Those independent variables exhibiting the highest degree of correlation were tested in combination with other explanatory variables. Selection of the final models (Table 2) was based on optimizing the dual criteria of explained variation (coefficient of correlation) and the inclusion of logical independent variables. A discussion of each of the models is presented in the following sections.

Home-Based Work Trips

The 2 independent variables that are included in the captive transit work trip production model pertain only to the resident population. The size of the resident labor force within each zone is directly related to the number of captive transit trips produced, whereas, as might be expected, automobile ownership exhibits a negative effect on transit use. The estimating model for captive transit work trip attractions also includes 2 variables—total employment and number of hotel and motel units. Both variables appear to be logical, and a reasonable degree of correlation was evidenced.

Home-Based Shop Trips

The independent variables included in the captive transit shopping trip production model reflect the dual character of trip-makers in the area—resident and tourist. The equation indicates that the number of residential households and hotel and motel units within each zone serves to define the total amount of shopping trips produced, whereas the automobile ownership variables limit the proportion of total demand allocated to the captive transit mode.

The model developed for captive transit shopping trip attractions includes 2 measures of commercial activity—general merchandise employment and commercial employment. The correlation coefficient for this equation is somewhat low, but it is acceptable for planning purposes.

TABLE 2

MODAL-SPLIT MODELS—CAPTIVE TRANSIT TRIP-GENERATION EQUATIONS

Trip Category	Equation	R
Production		
Home-based work	0.24119 LABFOR - 0.08835 RESAUT	0.8240
Home-based shop	0.07734 RESDU - 0.01484 RESAUT + 0.15936 HMUNIT - 1.01886 TOUAUT	0.8429
Home-based social-recreational	0.05104 RESDU - 0.01324 RESAUT + 0.11107 HMUNIT - 0.60464 TOUAUT	0.7386
Home-based school	0.04479 RESPOP	0.8106
Home-based miscellaneous	0.06504 RESDU - 0.02220 RESAUT + 0.05607 HMUNIT - 0.37454 TOUAUT	0.7790
Non-home-based	0.03866 TOTPOP - 0.04400 TOTAUT	0.7519
Attraction		
Home-based work	0.06725 HMUNIT + 0.07019 TOTEMP	0.7751
Home-based shop	0.37704 GMEMP + 0.07525 COMEMP	0.6206
Home-based social-recreational	0.00730 RESDU + 0.04891 HMUNIT + 0.07711 ED + AR	0.6669
Home-based school	0.37470 ENJRHS + 0.51166 ENSRHS + 0.17347 ENOTH + 0.04729 ENCOLL	0.8240
Home-based Miscellaneous	0.01344 HMUNIT + 0.05057 COMEMP	0.6694
Non-home-based	5.19916 + 0.20525 HBATTR	0.6844

Home-Based Social-Recreational Trips

The direct trip generation approach allows this particular trip purpose to be quantified separately rather than aggregated with other trip purposes as was done in the earlier MUATS model development work. This is especially important in Dade County where tourists, who might be expected to make a substantial proportion of social and recreational trips, represent a sizable minority of the winter season population.

The equation for social and recreational captive transit trip productions includes the same independent variables as the shopping trip production model—resident dwelling units, hotel and motel units, and resident and tourist automobile ownership rates.

The final estimating equation for social and recreational captive transit trip attractions includes 3 variables. The first, residential dwelling units, provides a measure of the visits made by friends and relatives. The second variable, hotel and motel units, is directly related to the social and recreational activities of tourists, and the third variable—eating, drinking, amusement, and recreational employment—is a measure of the social and recreational activity within a zone.

Home-Based School Trips

The equation that estimates school transit trip productions is similar to the work-trip model in that it is applicable only to the residents of the area. Initial model development attempts that included an automobile ownership variable were not successful, and it was necessary to define the relationship with only 1 variable—resident population. The failure of the automobile ownership variable to enter the equation can be attributed to the relatively young age of the school trip-makers. Junior and senior high school students, who constitute the overwhelming majority of school travelers, are by virtue of their age unable to drive and are, therefore, unaffected by the number of automobiles available to a household.

As might be expected, 4 school enrollment variables—junior high, senior high, college, and other—are included in the school transit trip attraction estimating equation. The first 2 variables, junior and senior high school enrollments, account for the great majority of school transit trip attractions, but the remaining 2 enrollment categories—college and other—are statistically significant and are included to complete the spectrum of the student trip-making population.

Home-Based Miscellaneous Trips

The miscellaneous captive transit trip production model includes the same 4 variables as the shop and social and recreational trip production models—resident dwelling units, hotel and motel units, and resident and tourist automobile ownership rates. This appears to be logical, and the resultant degree of correlation and the signs of the regression coefficients are also acceptable.

The inclusion of hotel and motel units and commercial employment in the miscellaneous captive transit trip-attraction model results in an equation that is responsive to logical measures of miscellaneous trip activity.

Non-Home-Based Trips

As the term implies, this trip category includes all travel within the study area in which neither trip end (production or attraction) is the home of the trip-maker. Numerous combinations of variables were considered during the development of the captive transit trip-production model, but most were rejected because of the low amount of explained variation. The final selected equation is similar to the other trip production formulas in that a demographic measure (in this case, total population) is directly proportional to the dependent variable whereas total automobile ownership is inversely proportional. Unsuccessful attempts were made to stratify population and automobile ownership by residents and tourists to obtain separate regression coefficients. Finally, the variables were aggregated for use in the model.

An investigation of the variables influencing the non-home-based captive transit trip attractions indicated that they were the same as those used in the other home-based

trip attraction models. A similar conclusion was reached by MUATS in the calibration of the total person trip model in this trip-purpose category. In view of this, the final model is directly related to the sum of all other home-based captive transit trip attractions in each zone.

Statistical Analyses of Captive Transit Models

Although the coefficient of correlation is an important statistical measure that provides an initial valuation of the model's acceptability, several other tests were also performed to evaluate the various captive transit trip-estimating equations. One key statistical test procedure is the F-ratio test that determines whether the model is a significant fit of the data. The F-test compares the variability about the fitted independent variables (numerator in F-ratio) with the inherent variability of the dependent variable (denominator). A computed value greater than the tabulated F-ratio for given conditions—degrees of freedom and level of confidence—indicates a significant fit. At the 95 percent confidence level, the computed F-ratio exceeded the tabular value for all captive transit models, denoting significant fits, as given in Table 3.

Another statistical measure of the trip-estimating equations is the t-test on the regression coefficients. This statistic is used to accept or reject the null hypothesis that a regression coefficient is equal to 0. The implication of the 0 condition is that the independent variable contributes nothing to the equation; i.e., any quantity times 0 equals 0. The results of this investigation (Table 3) indicated that the coefficient of each independent variable in all of the models is significant.

The final testing of the models consisted of evaluating the equations under base year conditions and comparing these results with the total number of actual transit trips. As given in Table 4, all trip models except the home-based social-recreational models slightly underestimate captive transit travel. The social-recreational trip models overestimate captive transit trips by 14.9 percent, but this error is insignificant when viewed in relation to the total pool of almost 174,000 trips. The models, with an overall predictive discrepancy of less than 3 percent in terms of total transit trips, are accurate simulators of existing captive transit travel.

CHOICE TRANSIT MODAL-SPLIT MODELS

The choice transit rider is the person who has an automobile available for his trip and can make the trip by automobile but for some reason or combination of reasons decides to use transit. The choice rider differs from the captive rider in that the relative attractiveness of the 2 alternate travel modes—transit and highway—has a significant effect on mode choice. In view of this, it is obvious that any models developed to forecast choice transit ridership should include some measure of the relative performance of the respective transportation systems (e.g., automobile travel time versus transit travel time).

The trip interchange modal-split technique was used in the development of the choice transit use models because this procedure is more responsive to specific improvements in the transportation system. Of interdistrict total person-trips, that proportion that is choice transit oriented is the dependent variable in this modal-split procedure. As a first step in the model-development phase, stepwise linear regression was performed on the district interchange data bank. The results of this initial computer analysis indicated that the number of trip categories should be consolidated. The home-based work and shop trip categories were kept separate, but all other nonschool purposes (including non-home-based) were combined. Because the number of choice school transit trips totaled less than 1 percent of the captive school transit trips, this trip-purpose category was excluded from further analysis.

Additional statistical investigation revealed that the work-trip model should include the travel time ratio (transit travel time divided by automobile travel time) and total employment density. Similarly, the shopping modal split was best explained by the travel-time ratio and commercial employment density. The last model, an aggregation of home-based social-recreational, miscellaneous, and non-home-based trip purposes correlated best with the travel-time ratio, total employment density, and the ratio of

TABLE 3

STATISTICAL ANALYSIS—CAPTIVE TRANSIT TRIP-GENERATION EQUATIONS

Trip Category	Variable	T-Test		F-Test	
		Computed	Tabular[a,b]	Computed	Tabular[a]
Production					
Home-based work	LABFOR	19.23	1.96	475.94	3.02
	RESAUT	10.12	1.96		
Home-based shop	RESDU	9.39	1.96	274.79	2.39
	RESAUT	3.76	1.96		
	HMUNIT	10.34	1.96		
	TOUAUT	5.49	1.96		
Home-based social-recreational	RESDU	5.77	1.96	134.42	2.39
	RESAUT	3.12	1.96		
	HMUNIT	6.71	1.96		
	TOUAUT	3.03	1.96		
Home-based miscellaneous	RESDU	9.91	1.96	172.85	2.39
	RESAUT	7.05	1.96		
	HMUNIT	4.56	1.96		
	TOUAUT	2.53	1.96		
Home-based school	RESPOP	29.39	1.96	863.88	3.86
Non-home-based	TOTPOP	16.86	1.96	292.74	3.02
	TOTAUT	7.16	1.96		
Attraction					
Home-based work	HMUNIT	6.82	1.96	338.54	3.02
	TOTEMP	20.09	1.96		
Home-based shop	GMEMP	10.51	1.96	140.92	3.02
	COMEMP	8.51	1.96		
Home-based social-recreational	RESDU	3.11	1.96	119.85	2.62
	HMUNIT	10.45	1.96		
	ED + AR	3.27	1.96		
Home-based miscellaneous	HMUNIT	3.78	1.96	182.69	3.02
	COMEMP	16.65	1.96		
Home-based school	ENJRHS	13.82	1.96	236.90	2.39
	ENSRHS	24.47	1.96		
	ENOTH	2.85	1.96		
	ENCOLL	6.17	1.96		
Non-home-based	HBATTR	19.91	1.96	396.45	3.86

[a]All tabular values are based on a 95 percent confidence interval.

[b]For large degrees of freedom, the t-statistic is very nearly the same as the Z or standard normal distribution.

total persons to total dwelling units. Even though the included variables in each of the respective equations appeared logical in terms of explaining the selection of travel mode, the degree of correlation was low. Further study of the degree of correlation of travel-time ratio with percentage of transit use indicated that a set of competitive diversion curves would be applicable in the Dade County area. The other explanatory variables identified in the respective equations would be included as parameters to the diversion concept. Because the diversion curve technique generally leads to a curvilinear relationship, linear regression was abandoned.

TABLE 4

COMPARISON OF ACTUAL AND ESTIMATED CAPTIVE TRANSIT TRIPS

Trip Category	Actual	Estimated	Difference (percent)
Home-based work	45,854	45,452	-0.9
Home-based shop	14,998	14,955	-0.3
Home-based social-recreational	8,606	9,893	+14.9
Home-based school	60,490	56,010	-7.4
Home-based miscellaneous	12,460	12,208	-2.0
Non-home-based	31,454	30,916	-1.7
Total	173,862	169,434	-2.6

This type of nonlinear approach has been commonly referred to as curve fitting because the data are plotted and curves are drawn through the points to reflect the modal-split relationship. Several regional transportation studies have used this method when linear relationships were not believed valid or could not be quantified from the data. This approach has an additional advantage in that it allows the planner to examine the data carefully and devise various stratifications of the source information.

The curve-fitting approach is iterative in that initial sets of curves are developed and tested for their ability to replicate existing transit travel. The curves are revised on the basis of the differences between actual and estimated transit travel and subsequently retested. The process of curve fitting and testing is repeated until accurate curves or models are obtained.

As shown in Figure 2, the final selected curvilinear models relate the travel-time ratio (transit versus automobile) to the proportion of total interdistrict travel demand that is choice transit travel. The curves for all 3 trip-purpose categories are similar in that they are curvilinear, and the travel-time ratio is inversely proportional to the percentage of transit-oriented travel. The work and shopping trip curves are stratified by the level of total employment density and commercial employment density respectively. The model for the other trip-purpose category conformed to the general shape of the curves for the work and shopping trip models. For this model, the data were stratified further by inclusion of a production end variable—total persons per total dwelling units.

TABLE 5

COMPARISON OF ACTUAL AND ESTIMATED CHOICE TRANSIT TRIPS

Trip Category	Coefficient Correlation	Actual	Estimated	Difference (percent)
Work	0.7752	6,014	6,529	+8.6
Shop	0.8521	2,274	2,160	-5.0
Other[a]	0.7785	7,254	7,072	-2.5
School	-	444	-	-
Total		15,986	15,761	-1.4

[a]Aggregation of home-based social-recreational, miscellaneous, and nonhome-based.

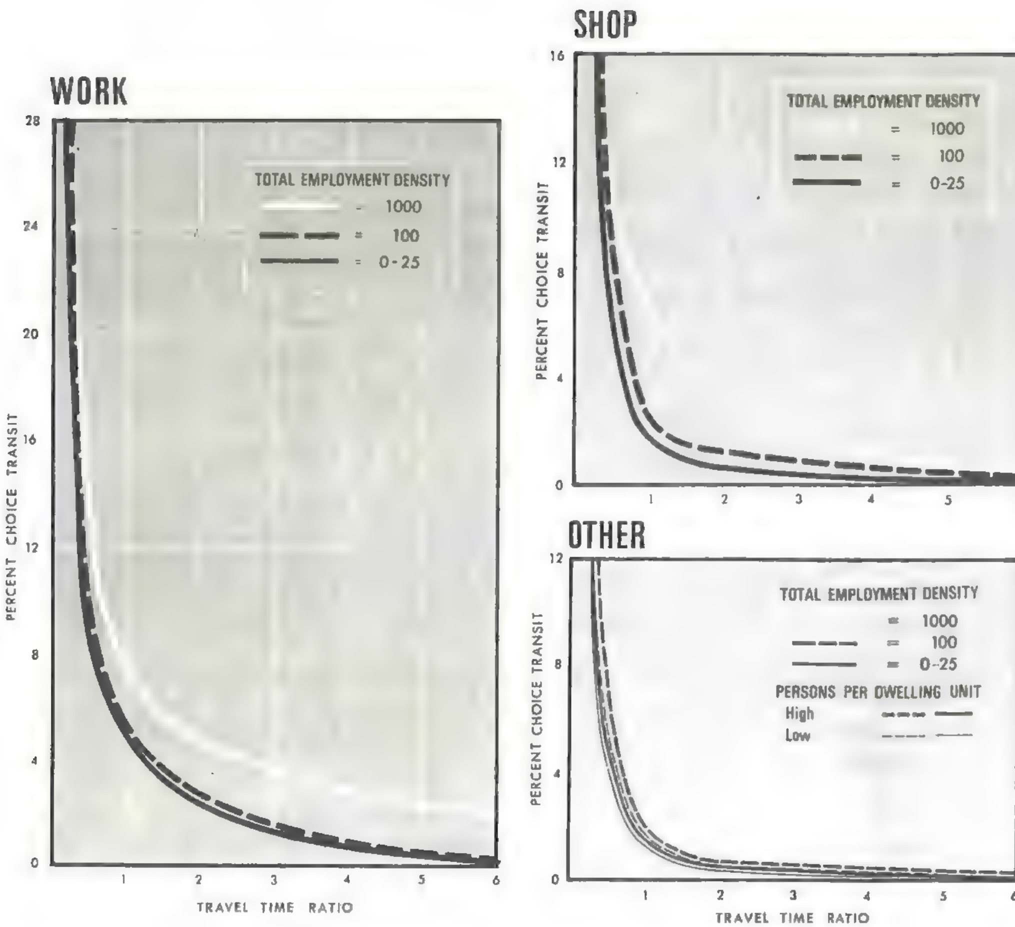


Figure 2. Choice transit modal-split curves.

Although the models appear logical and seem to explain the true causal relationship between transit use and the various travel-related variables, it is also of prime importance that they be able to simulate actual base year (1969) choice transit travel. One important measure of this ability is the coefficient of correlation that measures the amount of explained variation. This statistical term was computed by comparing the actual choice transit trips (1969) to the totals obtained by applying the curves to the respective 1969 total person-trip matrices on the basis of existing socioeconomic, land use, and transportation system conditions. As given in Table 5, the coefficients of correlation indicate a statistically good and significant relationship between the models and the real world. Table 5 also gives data showing that the discrepancy between actual and estimated choice transit trip totals ranges from a low of -2.5 percent for the other trips to a high of +8.6 percent for work trips. The overestimation of work trips almost balances the underestimation of the 2 remaining trip-purpose categories with a resulting error of less than 1-1/2 trips per hundred.

Another test of the models involved a comparison of the trip-length distribution for actual and estimated conditions. As shown in Figure 3, the trip-length distributions for all 3 transit trip purposes are very nearly the same. In addition, the differences between average trip lengths—actual and estimated—are within reasonable limits for all 3 trip-purpose categories. The largest discrepancy (4.7 percent) is found in the

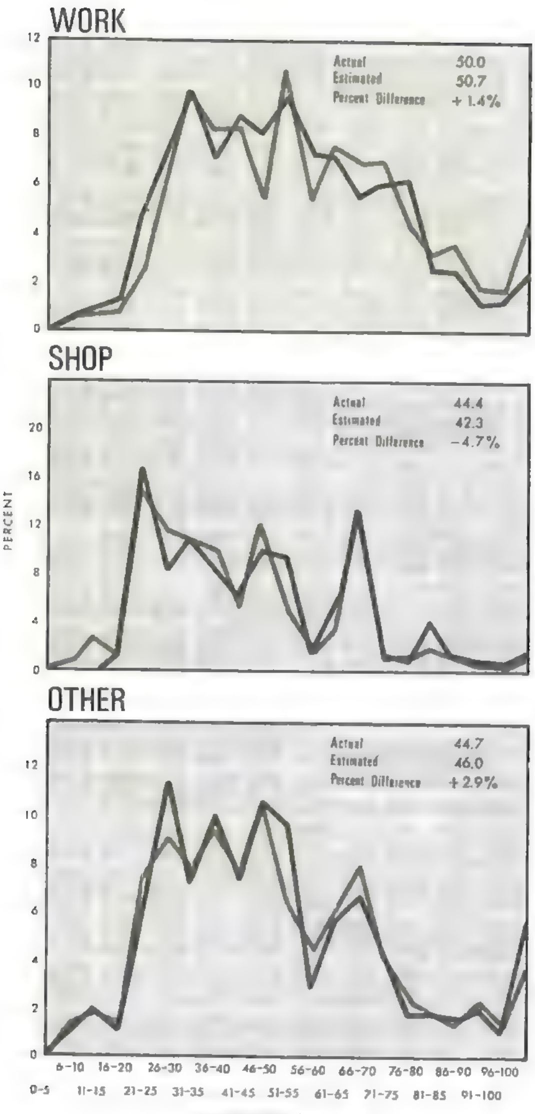


Figure 3. Choice transit trip-length distribution—actual versus estimated.

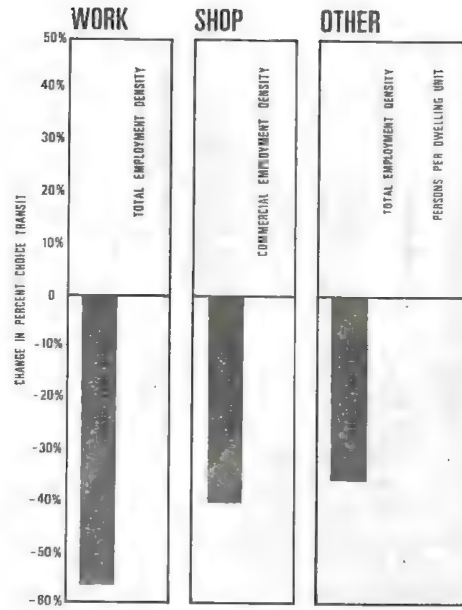


Figure 4. Reaction of percentage of choice transit to a 50 percent increase in travel-related variables.

shopping trip category, which also has the shortest average trip length. On the other hand, the work trip category has the smallest difference (estimated versus actual) and the longest average trip length. In general, it appears that all models adequately reproduce existing transit travel patterns as measured by trip length.

The previously discussed tests examined the ability of the models to simulate existing conditions. The assumption throughout is that, if the relationship quantified for 1 set of environmental conditions (1969) is valid, it will hold true for some future point in time—1985, for instance. Of prime importance in this particular analysis is how well the models respond to changes in study area conditions—specifically the transportation system. It was the failure of the 3 previous modal-split models in this area that prompted the investigation of the possibility of developing 2 sets of models.

One method used to gage model sensitivity to changes in each of the independent variables is to increase (or decrease) the value of each of the variables 50 percent above (or below) existing levels and then measure the impact on transit use (as defined by the percentage choice transit is of total person travel demand). Figure 4 shows that the travel-time ratio (transit or automobile) does indeed have a significant impact on percentage of choice transit trips. For example, a 50 percent increase in the travel-time ratio results in a 57 percent decrease in the percentage of choice work transit trips whereas, on the other hand, a 50 percent increase in employment density yields

only a 10 percent increase in the proportion of work transit trips. Similarly, the shopping trip model is also influenced the most by the transportation system variable (i.e., the travel-time ratio), whereas commercial employment density plays a lesser although still significant role in determining the modal split. A sensitivity analysis of the other trip-purpose model shows that it is about equally sensitive to the travel-time ratio and the persons per dwelling-unit variables. The impact of changes in the employment density variable in this model is not so great as that of the other 2 variables, but it is significant. This analysis indicated that models were responsive to changes in the included independent variables.

CONCLUSION

Various tests and analyses indicated that both the captive and choice modal-split models are reasonable and logical in terms of the included independent variables and statistically valid and capable of accurately simulating existing transit travel patterns. In addition, the choice models are responsive to changes in the transportation system.

The procedure of developing direct-generation equations for captive riders and trip-interchange models for choice riders recognizes the extremely different behavior patterns of the 2 groups in allocating their travel demand to the bimodal transportation system—the supply of transportation services. Segregation of the transit trips prior to model development permits the distinct characteristics of each transit group to emerge in their respective models.

This is especially important in urban areas that are currently served by buses traveling over congested city streets and that may have plans for high-speed rapid transit systems. Modal-split models in which transit riders are treated as a single group are calibrated from a transit trip-making population that is predominantly captive. It is questionable whether the modal-split relationships will be valid in future years for which major transit improvements are to be tested. Evaluation of the captive-choice formulas for Dade County with a comprehensive network of rapid transit facilities produced a shift in the dichotomy of transit patrons from only 1 choice rider in every 8 transit patrons in 1969 to 1 in 3 in 1985.

An additional benefit of the dual models is that 2 inventories of transit travel are available for study in future years. The travel movements of captive riders indicate where service must be supplied to accommodate the aspirations of the captive riders with regard to work, shopping, and social-recreational needs. This is especially important today when planners are more aware than ever before of the impact of mobility and its social consequences on urban residents. On the other hand, detailed analysis of the choice trips identifies those corridors where transit can play the most effective role in taking people from their cars and reducing the congestion on an already overburdened highway system.

The development of the 2 travel models—direct generation for captive riders and trip interchange for choice riders—appears to be a new and useful technique in evaluating transportation alternates of the future. The exploration of this approach in other cities would provide a useful data bank to evaluate more fully this approach and its applicability to other urban areas.

REFERENCES

1. Transit Usage Forecasting Techniques: A Review and New Directions. CONSAD Research Corporation. Pittsburgh. Final report prepared for New Systems Study Project, Urban Transportation Administration, U.S. Department of Housing and Urban Development, April 1968.
2. Fertal, M. J., Weiner, E., Balek, A. J., Sevin, A. F. Modal Split—Documentation of Nine Methods for Estimating Transit Usage. Bureau of Public Roads, U.S. Department of Commerce, Dec. 1966.
3. Development and Testing of Modal Split Models. Miami Urban Area Transportation Study, Tech. Rept. 4, July 1968.
4. Weiner, E. Modal Split Revisited. Traffic Quarterly, Jan. 1969, pp. 5-28.

DISAGGREGATE STOCHASTIC MODELS OF TRAVEL-MODE CHOICE

Shalom Reichman*, Department of Geography, The Hebrew University, Jerusalem, Israel; and
Peter R. Stopher**, Department of Environmental Engineering, Cornell University

This paper is concerned primarily with identifying the research and development needed to build general models of mode choice that are based on the modeling of individual behavior. The existing aggregate associative models are discussed in general terms, and their shortcomings are identified. The rationale of disaggregate behavioral models is then put forward, and the potentials of such a modeling technique are discussed. In this discussion, the present state of development of behavioral disaggregate models is indicated. Based on present experience in building these models, a set of research and development tasks is identified as being needed to develop more general operational models of this type. As each task is identified, suggested means of researching the problem are put forward.

•THE ESTIMATION of travel demand is a major part of most urban transportation studies. Because of the increasing complexity of deciding on investment priorities among transportation alternatives and between transportation and other urban and regional concerns, the accurate estimation and prediction of travel demand are becoming increasingly important as aids to the necessary decision-making process. The ability to predict travel demand more accurately is required both for problems within urban areas and for problems in major regional corridors. The existing travel-demand models, which have been developed largely for predicting intraurban travel, cannot meet the accuracy requirements of the decision-makers and policy-makers. As well as being too inaccurate for useful predictions of intraurban travel, these models are completely inadequate for use in predicting interurban travel. As a result of the shortcomings of existing models, there have been numerous attempts during the past few years to develop new strategies and techniques for modeling travel demand.

Within the problem area of travel demand, the specific problem of travel-mode choice, or modal split, is of considerable interest. Almost every transportation study has developed its own modal-split model, whereas models of the remaining elements of travel demand—trip generation, trip distribution, and assignment—are much more standardized. Mode-choice modeling derives a substantial amount of its added interest from the fact that it has considerable potentials in aiding investment decisions among transport modes and can potentially indicate the likely outcome of various decisions affecting today's ailing public transport undertakings. Although numerous modal-choice models have been developed, these models can largely be classified into a small number of categories wherein the properties of the individual models are susceptible to general description.

Recent work in the area of mode choice has produced a number of models that are concerned with determining the probabilities of choices of individuals. These models

*Mr. Reichman was visiting assistant professor in the Department of Geography and The Transportation Center at Northwestern University when this research was performed.

**Mr. Stopher was with the Department of Civil Engineering and The Transportation Center at Northwestern University when this research was performed.

Sponsored by Task Force on Value of Travel Time.

are based on the concept of applying theories, which concern human behavior and choice, to a disaggregate model structure for mode choice. For the most part, these models have been constructed for the work journey into the CBD of a large urban area. In general terms, this technique holds out considerable promise for the more accurate estimation and future prediction of mode choice.

It appears that this technique of modeling travel-mode choice has some considerable potential not only for more accurately estimating and forecasting modal shares but also for increasing basic understanding of the decision-making process involved. Among the additional benefits to be gained from this modeling technique is the ability to measure the comparative evaluation, in the mind of the traveler, of different system attributes. This clearly has a major use in making decisions among transport alternatives, particularly where patronage of a mode is of prime importance (1, 2).

The major concern of this paper is to describe and discuss the major areas of research and development that are currently indicated as being necessary for the realization of the potential of the disaggregate behavioral approach to modal-choice modeling. The first part of this paper is concerned with a classification and description of existing aggregate modal-choice models and with the identification of the major problems connected with these models. The second part of the paper describes a number of disaggregate behavioral models of mode choice that have been developed and details the potentials of the technique. It should be possible to realize these potentials when these models are generalized. The third part of the paper identifies and discusses the various problems that need to be solved if fully operational models of this type are built. These problems concern the identification of the best system, the user, and environmental characteristics for use in such models; the appropriate values for system characteristics; and some problems resulting from the complex trip patterns that are commonly encountered in urban areas. An outline of the needed research for the solution of each problem is given.

REVIEW OF MODAL-CHOICE MODELING

Modal choice, or modal split, is one of 4 models in the conventional urban transportation planning (UTP) process. This set of 4 models, which is collectively called the UTP package, has been reviewed in a number of papers (2, pp. 196-197; 3, 4). The limitations of this package, as a total travel-demand prediction tool, have been dealt with elsewhere at some length (2, pp. 197-199; 5, 6). Apart from the problems that arise within the UTP package as an entity, modal-split models have a number of specific disadvantages, stemming largely from the evolution of modeling requirements.

Since the first major transportation studies were carried out in the mid-1950's, various attempts have been made to build models of the modal-choice process. These models can be broadly classified into 3 principal groups. The earliest models attempted to predict modal choice at an aggregate level by using characteristics of the aggregated areas, such as zones and districts (7, 8, 9, 10, 11, 12, 13). These characteristics constituted socioeconomic measures of the population of the aggregated areas and measures of attraction of these areas in terms of the activity levels of various land uses. Because the models did not contain any characteristics of the modes, they could not respond to changes in these characteristics. Furthermore, because the socioeconomic measures incorporated are generally increasing (e.g., income, car ownership, and level of education), predictions of future modal shares from these models suggest that transit will be used by a dwindling proportion of the population, irrespective of any changes in mode characteristics.

The second generation of models came into existence in the early 1960's and incorporated some measures of the transportation system but still at an aggregate level (14, 15, 16). At first the system characteristics were still intractable because they comprised the fitted function of trip distribution from the gravity model (17). Later models began to incorporate actual system characteristics, such as time, cost, and frequency. An excellent example of this type of model is the diversion-curve model developed by the Traffic Research Corporation (16). However, these second-generation models are all aggregate models and have drawbacks in that they can

neither be transferred geographically nor readily subdivided; they are extremely sensitive to zoning; and the measurements of trip distances, particularly for short trips, are inaccurate because of the aggregation under which all trips are designated as originating or terminating at a point, the centroid of a zone. Also, these models are deterministic in the sense that they yield modal volumes rather than probabilities of using a particular mode. Deterministic models of this type have wide confidence limits statistically so that they are limited in their usefulness for explaining or predicting.

PRESENT RESEARCH ON MODAL-CHOICE MODELS

In an attempt to overcome some of the problems mentioned in the previous section, recent research has followed 2 directions. The first approach is an attempt to combine trip generation and modal-choice decisions in 1 model (18, 19, 20). A major conceptual problem, in the models developed from this approach, relates to the difficulty of establishing the theoretical structure for a decision-making process or situation. This is compounded when a number of decisions have to be made, some of which are contingent on and some independent of the others. None of the models so far proposed has addressed itself specifically to the elaboration of this multiple and simultaneous decision-making process.

Although these models are conceived only as generation and modal-choice models, to make them operational, and hence to make them even more complex, requires that trip distribution be included. This is necessary because a combination of a modal-choice model and generation model is not defined operationally because the specific trip interchange must be known before values can be obtained for the system charactertistics operating on modal choice and generation.

Finally, several problems are encountered in the estimation processes. Considerable difficulty exists in formalizing mathematically the complex decision-making processes already described. Attempts at doing this have not been conspicuously successful. The statistical reliability of the combined generation and distribution modal-choice model remains to be firmly established. This is due to the complexity of the total choice mechanism and the relative simplicity required for operational purposes as well as the aggregation bias. It is not surprising, therefore, that in a discussion of the statistical reliability of one of these models, the following conclusion was reached (21):

> The high level of residual error in estimating the total choice mechanism (as opposed to a single aspect) should be regarded as a danger signal by the planner. The result implies high uncertainty in our predictions of the effects of changes in the transportation system. When account is taken of sampling errors and errors in predicting independent variables, in addition to the generally low correlation statistics, it is clear that the uncertainty in predicting future origin and destination traffic movements is very great indeed. The planner must therefore be extremely cautious in his decisions and explicitly recognize that his evaluations are subject to this uncertainty.

The second approach to building new models of travel-mode choice is based on the application of theories of individual behavior in making choices and yields, initially, a disaggregate probabilistic model. However, unlike the first approach, these disaggregate behavioral models are conceived of as operating within the same model structure as the conventional UTP package of models.

The main advantage of this approach is that modal choice is analagous to the typical marketplace decision situation (22). The individual buyer, namely the traveler, has to choose between a number of goods or services according to their attributes and his set of preferences. It is felt that the present state of behavioral modeling permits relatively sound operational formulation of this choice process.

This approach represents a major change in emphasis in 2 respects: The models are disaggregate, and they attempt to model individual behavior on the basis of mode and user attributes (23, 24, 25, 26). At present, such models have not been developed specifically as a stage in a transportation study and therefore do not fit well in the

current overall modeling processes used in such studies. They have been developed to predict individual behavior by assigning to individuals probabilities of using modes in a binary choice situation. It appears that these models can better reproduce existing conditions, regardless of geographic location, and give more reasonable responses to forecasts of mode and user characteristics than models of the type described in the earlier section.

PROPERTIES OF DISAGGREGATE STOCHASTIC MODELS

The models being developed within the second approach discussed in the preceding section are disaggregate stochastic models (2). The models are stochastic in that they predict the probability that an individual will make a specific choice. This probability is assigned on the basis of the consideration of user and system characteristics, and this procedure is most consistent with modern theories of human discrimination and choice (27, 28). These theories state that every human decision is, in essence, probabilistic. In addition to providing a basis in behavior to the concept of stochastic disaggregate models, these theories lead to 2 conclusions that are extremely important in formulating models of this type. First, disaggregate stochastic models of this type can be formulated with a relatively small number of variables required to achieve good predictions. Second, people do not have irrational or unquantifiable biases toward specific alternative choices.

Placed in the context of modal choice, this approach effectively states that an individual will choose a mode with a probability determined by trip considerations and his own scaling of the effectiveness of the alternatives for that trip purpose.

Thus, the stochastic model of modal choice may be considered as a translation of the theoretical elements of decision-making into operational terms. Based on the revealed preferences by travelers in their modal choices, the models distinguish between the attraction and disutility of the system characteristics and user characteristics that affect the preference scale of system attributes. More specifically, the models can incorporate a variety of system characteristics and, provided that adequate quantification can be achieved, may include attributes derived from attitudinal studies (29, 30, 31). These models have several properties that are of help in analyzing modal choice.

First, these models have greater predictive validity than conventional models because they are based on the behavior patterns of individuals rather than on statistically derived correlations in aggregate analysis.

A second property of these models is that they are based on the smallest element of the population—the individual. As such, the large variances attributed to problems of zoning or of aggregation of population attributes are eliminated (5, 32, 33, 34, 35). Thus, the predictions from the models must have much narrower confidence limits than those of aggregate models.

A third property, also stemming from the disaggregate approach, relates to the eventual aggregation that must be used for the models to be applied to large urban areas. Because the models are constructed from the smallest population elements, the level of aggregation can be determined from the models as the precision of measurement required to yield the desired predictive accuracy. In contrast to earlier modal-choice models, a stochastic disaggregate model is a deductive model in that it determines a priori the variables to be measured and the level of aggregation and the groupings for aggregation that are required for operation of the model.

A fourth property of stochastic disaggregate models is that they provide a basis for inferring the relative values that people place on various characteristics of the transportation systems. These values may be derived from an examination of the relative weights of each system characteristic on the choice process described by the model. The behavioral basis of the models requires that, by definition, the values so obtained be behaviorally consistent.

CURRENT STATE OF THE ART

The stochastic modal-choice models (see Appendix and 47), developed at present, have been built as binary-choice models by using a variety of statistical tools. These

tools are probit analysis, logit analysis, and discriminant analysis. These statistical techniques have been used because they seem, to some extent, to reflect the choice situation as it is perceived to exist by the transportation planner.

Four models have been developed by using the theory of discriminant analysis (24, 36, 37, 38). This theory hypothesizes that the total traveling population of a study area can be viewed as comprising two or more distinct subpopulations, divided on the basis of the modes they use. The task of discriminant analysis (39) in this case is to determine a function of user and system attributes that best discriminates between the subpopulations. All four of these models work from sets of binary choices; i.e., they identify a series of populations containing people with a binary choice and operate on these binary choices alone. Thus, the task is to determine a set of discriminant functions, D_{ij} (where D_{ij} is the discriminant function between mode i and mode j), that minimize the number of members of the population that will be misclassified (in terms of their choice of mode) by the model (Fig. 1).

Probit analysis (25, 40, 41) is a statistical technique originally derived in work on toxicology. It is based on the premise that, if members of a population are subjected to a stimulus that can range over an infinite scale, the frequency of responses to the stimulus (assuming that the response is a 0, 1 response, i.e., it either occurs or does not occur) will be normally distributed. Thus, if a cumulative plot of members of the population who have already responded is drawn against the value of the stimulus, the curve that results will be a normal sigmoid (S-shaped) curve (Fig. 2). This theory is applied to the modal-choice situation by assuming that the stimulus is made up of the relative disutilities of travel on 2 modes in a binary situation and of the characteristics of the user. The estimation problem becomes one of assigning coefficients, or weights, to the disutilities and user characteristics in the probit equation.

The last of the principal analytical techniques and one that has been used for this type of modeling is logit analysis (26, 42). In its simplest form, the linear logit model states mathematically that the probability of the occurrence of an event varies with respect to a function G(X) as a sigmoid curve called the logistic curve. This model may be written mathematically as

$$P_1 = \frac{1}{1 + e^{G(X)}}$$

which is identical to part of the discriminant analysis model, although it should be noted that the means of evaluating the function G(X) is very different in these 2 cases. This model is applied to modal choice by defining the event mentioned previously as being the choice of 1 mode in preference to the other.

It should be noted that, although probit and logit analysis are mathematically different and are based on dissimilar premises, operationally there is a considerable similarity in the results produced by the 2 techniques. Problems concerning which of the 3 techniques mentioned previously is most appropriate for the building of behavioral modal-choice models are discussed in more detail in the remainder of this paper.

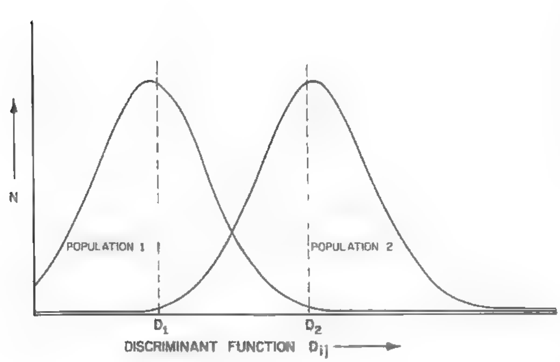


Figure 1. Frequency plot of values of the discriminant function in a binary population.

PROBLEM IDENTIFICATION

In the building of operational disaggregate stochastic models of modal choice, a number of problems are encountered. This section of the paper is concerned with identifying a number of these problems and with proposing

possible strategies for solving the problems. Six problem areas are identified as requiring research to develop useful, operational models of modal choice.

The first problem area concerns the parameters needed to describe the transportation systems among which the traveler must make his choice. In particular, the inclusion of some measures, which represent comfort, convenience, and safety, seem desirable but present many problems in the framework of a quantitative model structure. The second problem area concerns the set of user characteristics to be included in the model and the form in which such characteristics should operate in the mathematical function.

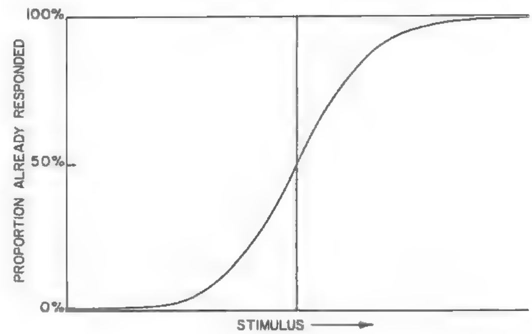


Figure 2. Cumulative plot of response to a varying stimulus.

Third, the various coefficients of the model are likely to depend on the value to be derived from the trip. The determination of a good set of trip classifications related to trip value is clearly needed for the development of these models of modal choice.

Fourth, frequent questions have been raised concerning what the appropriate values are for the system parameters—those perceived by the traveler, or those measured by an independent observer. Again, this represents a major problem area in developing disaggregate behavioral models.

The fifth problem area concerns temporal and geographic, or spatial, transferability. To develop models that are not restricted in application to specific locations and times of day requires an examination of possible dependencies between coefficient values and the dimensions of time and space and a determination of the means to handle such dependencies that are found to exist. Finally, problems are encountered in determining the ways to handle trip segmentation and a multiple-choice environment. Both of these problems are directly related to the mathematical structure of the model.

These are, in outline, the central problems that must be resolved to achieve operational disaggregate behavioral models of modal choice. Each of these problems is discussed in detail, and possible strategies for solution are indicated in the following sections.

DEVELOPMENT OF A BEHAVIORAL MODAL-CHOICE MODEL

Identification of System Variables

System variables can be classified as either exogenous or endogenous to the decision-maker. The exogenous system variables may include overall travel times, overall travel costs, and travel time reliability. This last refers to all those factors that intervene by adding delays or uncertainty in achieving trip goals. These would include segments of delay and uncertainty such as walking, waiting, transfer, and a built-in unreliability. This unreliability includes a component learned by the traveler from experience with schedule variance or traffic congestion and a component of unpredicted delays arising from system accidents (e.g., vehicle breakdown and weather).

It may be hypothesized that a detailed analysis of system unreliability is equivalent in fact to what is commonly known as convenience. Convenience is understood here as an exogenous property of the system reflecting the extent of impediments to travel and can be measured by the delays that the system imposes on the users and that increase the uncertainty or perceived unreliability of the system.

Safety is not considered to be an independent variable for modeling purposes. As far as it is an objective system attribute, it is already represented in the evaluation of the unpredictability referred to previously. Insofar as it is a projection of the individual

anxiety of the traveler onto system characteristics, little additional explanatory power will be achieved by incorporating it.

The exogenous system variables can be treated in several ways with respect to the method of inclusion in a disaggregate behavioral modal-choice model. Most of the models developed in the past have used either ratios of costs and of times or differences of costs and of times. The use of ratios in a behavioral model implies that the decision-maker views time or cost differences on a relative scale, i.e. his decision would be affected to the same extent by a choice of travel times of 10 and 5 min as by a choice of 60 and 30 min. The use of differences, however, implies an absolute valuation of times and costs so that the choice between travel times of 10 and 5 min would now be equivalent to a choice between 60 and 55 min. A third alternative, which has not been examined in detail before, is to use differences relative to the total overall cost or time. This would imply that the choice of the decision-maker is influenced by the absolute value of the cost or time saving relative to the total outlay of cost or time that he must bear if he is going to make the trip. It is clear that this method of inclusion would give rise to significantly different values of the time and cost components of trips, particularly in the case of discretionary trips and of interurban movements where choice is between, for example, air and ground common carriers. Each of these alternatives may be examined both by scaling experiments and by their relative performance in each model in predicting modal choice.

The main endogenous system variable is comfort. This may be treated by carrying out a cross comparison among mode alternatives of their comfort characteristics. The characteristics to be examined would include the space for each passenger, shock and vibration, noise, privacy, and other variables relating to the physiological and psychological characteristics of movement systems. On the basis of available data on acceptable values of these characteristics to users, an interval scale might be developed that would allow measurement of differences among modes in terms of comfort. In addition, an attempt might be made to develop a psychometric ratio scale that, if successful, would generate a single value of comfort that could be used to evaluate any existing mode or a new technology. It should be emphasized that the weighting procedure, referred to earlier, of the comfort variables is fundamentally different from the treatment of comfort as a residual variable.

User Characteristics

When the behavior of individuals is considered in decision-making, it is desirable to consider their individual utility functions. Detailed consideration of these utility functions is not possible, however, so socioeconomic characteristics of the users are generally introduced instead. These characteristics serve as proxies to represent the average behavior of the individual decision-makers.

The user characteristics that are generally considered for inclusion in models of this type are income, age, sex, stage in the family life cycle, and car availability. For stochastic disaggregate models, these characteristics should refer, as far as possible, to the individual and not to the household. Experience of the use of these general categories of characteristics suggests that problems of collinearity may exist among several of the variables listed. This problem may be tackled by the use of techniques such as factor analysis, stepwise regression, and covariance analysis to attempt, as far as possible, to eliminate variables that do not add significantly to the explanatory power of the model.

Previous models of modal choice have largely introduced user characteristics as additive terms along with the system characteristics. This form of inclusion is effectively a behavioral assumption that the choice is made on the basis of system characteristics by themselves, with the addition of an individual bias. This bias is the additive value of the user characteristics included in the model. An alternative assumption appears to be more consistent with behavior theory; that is, the weights (coefficients) attached to each of the system characteristics are dependent on the individual (i.e., the user characteristics). In other words, this assumption implies that an individual bias exists on the importance of each system characteristic rather than on the final choice.

This form of inclusion can be carried out by subdividing the population, on which calibration is performed, into a number of classes representing ranges of each of the user characteristics. Models are then calibrated for each class of the population, and associative relationships are sought between variations in the coefficients of each system attribute and the user attribute values of each class. The extent to which subdivision of the population is possible will inevitably depend on the data sources available and the use to which the model will be put.

Trip Characteristics

When the values attached by each decision-maker to the system atrributes are examined, another source of variation can be hypothesized. This variation relates to the fact that travel can be regarded as a derived demand (42) that is required to permit another activity to be carried out. As such, it seems reasonable to hypothesize that the values of system attributes will be related to the value to be derived by the tripmaker from the activity that he will carry out at the end of his trip. Conventionally, attempts are not made to determine the value derived from an activity, but a proxy—the trip purpose—is used.

In a definition of the trip purposes for disaggregate behavioral models, 2 constraints have to be considered. The first of these is the constraint imposed on the inclusion of trip purposes by the available data. Data available at present are largely work-trip oriented, and, therefore, there is a need for a widening of the data base in this respect. In addition to this, a further constraint is imposed by a consideration of the definition of trip purposes that are currently used in most studies (44). This is an area in which considerable research is needed.

In addition to these principal problems, there are several further problem areas for which some assumptions are necessary so that a viable model formulation can proceed. These problems are outlined in the next section.

Values of System Characteristics

A consideration important in the formulation of a behavioral model is the relation between objective and subjective estimations by the traveler of the system characteristics. The difference between true and perceived values of, say, travel time or costs arise from 2 sources. One is inadequate information about, or experience with, alternative modes. With inadequate information, people will make choices by filling in the necessary judgments on a subjective basis. Obviously, this may bear little relation to objective reality, but it is still the basis for modal choice. The second is a bias that persists even with knowledge of alternatives. By definition, this bias is a stable preference function.

It is obvious that, for predictive purposes, the former process is most critical because it may be assumed that any effects of a stable preference function may be resolved by a simple linear transformation. In fact, model calibration achieves this. The problem caused by lack of information is that a priori there is no way of knowing how these deviations from objectivity are distributed nor at what rate learning modifies the subjective values to make them approach the objective ones. Ideally, if the distribution of subjective values around objective values of the system characteristics are normal, the errors will sum to 0. Alternatively, a consistent relationship between subjective and objective values may be able to be determined. These effects should show up as unexplained variance in the model that, if it were high, would be the basis of a recommendation for extensive analysis in this area.

Environmental Characteristics

The environmental characteristics are those characteristics that are independent of the system and user attributes and can be broadly defined as the temporal and spatial dimensions within which the modal-choice decision is made. The temporal dimension can be typified by the time of day when the trip is made. The time of day can be seen to affect the modal-choice decision process in 3 ways: It will be a partial determinant of system characteristics because of variations in congestion or loading,

transit frequencies, and journey speeds and costs; it will also be a partial determinant of trip purpose; and probably the values of time savings in particular will vary with time of day in relation to the possible uses of such time savings.

Characteristics describing the spatial dimension appear to be necessary to allow the models to be transferable, in terms of their structure and parameters, over a range of urban areas. These characteristics would be constructed as central measures (i.e., 1 measure of each variable for an urban area) and would stand as proxies for the range of distances, costs, and the like that vary from urban area to urban area. Probable variables that would be used in this manner would be the size, residential density, and age of the city (12, 45).

The extent to which these characteristics can be incorporated in the models will depend on available data. Data restrictions that currently exist would not permit analysis of these problems because most of the available data refer to only 1 city. Similarly, the data restrictions to the morning work trip largely preclude detailed analysis of time of day at present. Again, however, there is a need for further data to assist analysis of this type.

Model Characteristics

Two basic problems need to be tackled in determining the specific form of the behavioral modal choice model. The first of these concerns the structure of trip-making. The majority of transit trips usually consists of 3 segments: access to the transit facility, line-haul, and egress to the final destination. For the automobile, the segments would be line-haul, parking, and egress to the final destination. The problem that has to be considered is how to treat these trip segments in the model. Several possible alternatives include a model based on the line-haul mode only (with average or best system characteristics assumed for collection and distribution), a model in which each possible combination of modes is treated as a separate mode, or an hierarchical set of models that determine the choices for each segment of the trip. These alternatives need to be examined in terms of the operational and conceptual requirements placed on the model.

The second problem concerns the range of choices to be considered in the models. Previous models have been developed as binary-choice models in which it is assumed that the trip-maker reduces his choice problem to a choice between 2 alternatives. The drawback to this approach is the difficulty of identifying which 2 modes represent the final choice of each individual. Alternatively, the models could be developed as multinomial models in which the choices between all possible modes available to the trip-maker are incorporated simultaneously. A theoretical model structure has been developed (42, 46, 48) for modeling a multinomial situation, but its properties, in terms of sensitivities in particular, have not been investigated.

Research is needed here in 2 areas. First, an investigation of the use of the multinomial model is required. Second, research is required to determine the problems involved in the collapse of an individual's choice to 2 alternatives.

Conclusions

Several conclusions can be drawn from the foregoing discussion concerning disaggregate, stochastic, behavioral models of travel-mode choice. First, it is clear that these models can offer several advantages over conventional aggregate deterministic models in estimating and forecasting modal shares. These advantages are principally that the use of choice theory, applied at an initially totally disaggregate level, will lead to models that are spatially and temporally more stable than conventional models. In particular, this leads to a much greater confidence in the resulting forecasts from the models, when used to indicate the likely consequences of particular policies and investment decisions.

Second, although the existing models of this type are restricted both in the type of trips and the variables considered for model calibration, the developmental problems that will be encountered in extending their applicability and usefulness appear to be generally straightforward and readily amenable to solution. In fact, the major problems

reside in data collection for tackling these various model developments, and past experience with these models suggests that the data problems are by no means insurmountable.

The third conclusion concerns the information that these models can yield concerning traveler-evaluation of system characteristics. The models already developed have been used to obtain monetary values for the value of travel time. As other system characteristics are explicitly incorporated in the models, other comparative values can also be determined. These values may have considerable use in assisting decision-making on alternative investments in public transport, among other uses. Here it may help to determine whether travelers would prefer more frequent service with higher fares or less frequent service and lower fares, whether they would prefer faster travel or air-conditioning, and so on.

Finally, the successful development of disaggregate, stochastic, behavioral models of modal choice will open up considerable potentials for developing models of the remainder of the travel-demand process on the same basis (2). Considerable understanding of the mechanism of traveler choices will accrue from the modal-choice models, and this understanding can readily be applied to other choice situations with a consequent increase in the accuracy and reliability of travel forecasts.

REFERENCES

1. Stopher, P. R. Value of Time from Modal Choice Analysis: Some Problems. Paper prepared for meeting of HRB Task Force on the Value of Time, January 1970.
2. Stopher, P. R., and Lisco, T. E. Modelling Travel Demand: A Disaggregate Behavioural Approach—Issues and Applications. Transportation Research Forum Proc., 1970, pp. 195-214.
3. Davis, H. E. Urban Transportation Planning—Introduction to Methodology. In Defining Transportation Requirements, ASME, 1969, pp. 9-23.
4. Martin, B. V., Memmott, F. W., and Bone, A. J. Principles and Techniques of Predicting Future Demand for Urban Area Transportation. Dept. of Civil Engineering, M.I.T., Res. Rept. 38, June 1961.
5. McCarthy, G. M. Multiple-Regression Analysis of Household Trip Generation—A Critique. Highway Research Record 297, 1969, pp. 31-43.
6. Deutschman, H. Urban Transportation Planning. Sources of Information on Urban Transportation, ASCE, Rept. 4, June 1968.
7. Adams, W. T. Factors Influencing Mass Transit and Automobile Travel in Urban Areas. Public Roads, Vol. 30, 1959, pp. 256-260.
8. Howe, J. J. Modal Split of CBD Trips. C.A.T.S. Research News, Chicago Area Transportation Study, Vol. 2, No. 12, 1958, pp. 3-10.
9. Biciunnas, A. E. Modal Split of First Work Trips—1956 and 1960. C.A.T.S. Research News, Chicago Area Transportation Study, Vol. 6, No. 1, 1964, pp. 12-13.
10. Sharkey, R. H. A Comparison of Modal Stratification of Trips by Distance From the CBD. C.A.T.S. Research News, Chicago Area Transportation Study, Vol. 2, No. 5, 1958, pp. 14-20.
11. Sharkey, R. H. The Effect of Land Use and Other Variables on Mass Transit Usage in the C.A.T.S. Area. C.A.T.S. Research News, Chicago Area Transportation Study, Vol. 3, No. 1, 1959, pp. 3-10.
12. Schnore, L. F. The Use of Public Transportation in Urban Areas. Traffic Quarterly, Vol. 16, 1962.
13. Hadden, J. K. The Use of Public Transportation in Milwaukee, Wisconsin. Traffic Quarterly, Vol. 18, 1964, pp. 219-232.
14. Quinby, H. D. Traffic Distribution Forecasts: Highway and Transit. Traffic Engineering, Vol. 31, No. 5, 1961.
15. Hamburg, J. R., and Guinn, C. R. A Modal-Choice Model: Description of Basic Concepts. New York State Department of Public Works, Dec. 1966.
16. Hill, D. M., and von Cube, H. G. Development of a Model for Forecasting Travel Mode Choice in Urban Areas. Highway Research Record 38, 1963, pp. 78-96.

17. Calibrating and Testing a Gravity Model for Any Size Urban Area. Office of Planning, U.S. Bureau of Public Roads, 1963.
18. Quandt, R. E., and Baumol. W. J. The Demand for Abstract Transport Modes: Theory and Measurements. Jour. of Regional Science, Vol. 6, No. 2, 1966, pp. 13-26.
19. Domencich, T. A., Kraft, G., and Valette, J. P. Estimation of Urban Passenger Travel Behavior: An Economic Demand Model. Highway Research Record 238, 1968, pp. 64-78.
20. McLynn, J. M., and Watkins, R. H. Multimode Assignment Model. Paper prepared for National Bureau of Standards, Washington, D.C., Aug. 1965.
21. Charles Rivers Associates, Inc. An Evaluation of Free Transit Service, NTIS, Springield, Va., PB 179 845, Aug. 1968, p. 52.
22. McLynn, J. M., Goldman, A. J., Meyers, P. R., and Watkins, R. H. Analysis of a Market-Split Model. Northeast Corridor Transportation Project, Tech. Paper 8, April 1967.
23. Warner, S. L. Stochastic Choice of Mode in Urban Travel: A Study in Binary Choice. Northwestern Univ. Press, 1962.
24. Quarmby, D. A. Choice of Travel Mode for the Journey to Work: Some Findings. Jour. of Transport Economic and Policy, Vol. 1, No. 3, 1967, pp. 273-314.
25. Lisco, T. E. The Value of Commuter's Travel Time: A Study in Urban Transportation. Dept. of Economics, Univ. of Chicago, PhD dissertation, 1967.
26. Stopher, P. R. A Probability Model of Travel Mode Choice for the Work Journey. Highway Research Record 283, 1969, pp. 57-65.
27. Bock, R. D., and Jones, L. V. The Measurement and Prediction of Judgement and Choice. Holden-Day, San Francisco, 1968.
28. Luce, R. D. Individual Choice Behavior. John Wiley and Sons, New York, 1959.
29. User Determined Attributes of Ideal Transportation Systems: An Empirical Study. Dept. of Business Administration, Univ. of Maryland, 1966.
30. Lansing, J. B., and Hendricks, G. Automobile Ownership and Residential Density. Survey Research Center, Inst. for Social Research, Univ. of Michigan, 1967.
31. McMillan, R. K., and Assael, H. National Survey of Transportation Attitudes and Behavior: Phase 1—Summary Report. NCHRP Rept. 49, 1968, 71 pp.
32. Worrall, R. D. Toward a Longitudinal Model of Household Travel. Northwestern Univ., PhD dissertation, 1966, Chap. 2, Section ii.
33. Goldman, L. A. Some Alternatives to Ecological Correlation. American Jour. of Sociology, Vol. 64, 1959, pp. 610-625.
34. Alker, H. R., Jr. A Typology of Ecological Fallacies. In Quantitative Ecological Analysis in the Social Sciences. (Dogen, M., and Rokkan, S., eds.), M.I.T. Press, 1969, pp. 69-86.
35. Robinson, W. S. Ecological Correlations and the Behaviour of Individuals. American Sociological Review, Vol. 15, 1950, pp. 351-357.
36. Mongini, A. Some Aspects of Discriminant Functions and Other Interurban Modal Split Models. Report submitted to Northeast Corridor Transportation Project, Oct. 1965.
37. McGillivray, R. G. Binary Choice of Transport Mode in the San Francisco Bay Area. Dept. of Economics, Univ. of California, PhD dissertation, 1969.
38. Bock, F. C. Factors Influencing Modal Trip Assignment. NCHRP Rept. 57, 1968, 78 pp.
39. Fisher, R. A. The Use of Multiple Measurements in Taxonomic Problems. Ann. Eugen., London, Vol. 7, 1936.
40. Finney, D. J. Probit Analysis. Cambridge Univ. Press, 1964.
41. Lave, C. A. Modal Choice in Urban Transportation: A Behavioral Approach. Dept. of Economics, Stanford Univ., PhD dissertation, 1968.
42. Thiel, H. A Multinomial Extension of the Linear Logit Model. Rept. 6631, Nov. 1966.
43. Oi, W. Y., and Shuldiner, P. W. An Analysis of Urban Travel Demands. Northwestern Univ. Press, 1962, Chap. 2.

44. Manual of Procedures for a Home Interview Traffic Study. Bureau of Public Roads, U.S. Department of Commerce, Oct. 1954.
45. Aaugeenbrug, R. T. Automobile Commuting: A Geographic Analysis of Private Car Use in the Daily Journey to Work in Large Cities. Dept. of Geography, Univ. of Wisconsin, PhD dissertation, 1965, Chap. 7.
46. Stopher, P. R. A Multinomial Extension of the Binary Logit Model for Choice of Mode of Travel. Northwestern Univ., unpublished paper, 1969.
47. McGillivray, R. G. Demand and Choice Models of Modal Split. Jour. Trans. Econ. and Policy, Vol. 4, No. 2, 1970, pp. 192-207.
48. Rassam, P. R., Ellis, R. H., and Bennett, J. C. The N-Dimensional Logit Model: Development and Application. Paper presented at HRB 50th Annual Meeting, and included in this Record.

APPENDIX

CURRENT STATE OF THE ART

A number of models of modal choice that have been developed, recently, are based on the rationale of disaggregate, behavioral modeling. These models are summarized here. The models may be classified according to the mathematical technique used, i.e., discriminant, probit, or logit analysis.

The 4 discriminant models, referred to in the paper were developed by Quarmby, Mongini, McGillivray, and the Illinois Institute of Technology Research Institute. Quarmby (24) and McGillivray (37) each propose a discriminant function that effectively comprises system characteristics and user characteristics, although Quarmby used differences in system characteristics and McGillivray used ratios of system characteristics. Both models added user characteristics as linear terms in the discriminant function. Mongini (36) used just time differences in his pilot study but suggested that other elements of the total disutility of travel should be incorporated. IIT Research Institute (38) used a series of associative and correlation tests to determine variables to be included, and this resulted in a total of 39 variables describing user and system characteristics in various ways.

The use of a discriminant analysis results in the formulation of a discriminant function that, for the models developed by Quarmby, McGillivray, Mongini, and IIT Research Institute, may be generalized as

$$D_{ij} = F(X_{ki}, X_{kj}) + G(Y_m)$$

where X_{ki}, X_{kj} are the values of the kth attribute of modes i and j respectively; Y_m are the user attributes; $G(Y_m)$ is a function $\sum_{m=1}^{h} \beta_m Y_m$; $F(X_{ki}, X_{kj})$ is either a function $F(X_{ki}, X_{kj}) = \sum_{k=1}^{n} \alpha_k (X_{ki} - K_{kj})$ or a function $F(X_{ki}, X_{kj}) = \sum_{k=1}^{n} \alpha_k (X_{ki}/X_{kj})$.

The model is used as follows: If D_{ij} is less than a value D_1 (Fig. 1), the individual is classified as a user of mode 1; and, if D_{ij} is greater than a value D_2, he is classified as a user of mode 2. If D_{ij} lies between D_1 and D_2, the probability that the individual is a user of mode 1 can be shown to be

$$P = \frac{1}{1 + e^{D_{ij}}}$$

so that the resulting modal-choice model could be written as

$$P_1 = \begin{cases} 1 & D_{ij} < D_1 \\ 1/(1 + e)^{D_{ij}} & D_1 < D_{ij} < D_2 \\ 0 & D_{ij} > D_2 \end{cases}$$

where P_1 is the probability that an individual will choose mode 1.

Both Lisco (25) and Lave (41) used probit analysis as the basis of the calibration of of modal-choice model. Using this theory, Lisco and Lave each proposed a probit equation of the following form:

$$Y = \sum_{k=1}^{m} [\alpha_k (X_{1k} - X_{2k})] + \sum_{\ell=1}^{n} [\beta_\ell (Y_\ell)]$$

where the X's and Y's are as defined for discriminant analysis.

The system characteristics in both cases include travel costs and travel times, and the user characteristics include family size, income, sex, and age in Lisco's model and sex and age in Lave's model. Lave also attempted to include a measure of comfort in his model.

The probit equation is used for modal choice in the following manner: The value of the probit, Y, represents the number of the standard deviations away from the mean of a normal distribution. If standard statistical tables of the cumulative normal distribution are used, the probability, corresponding to this number of standard deviations from the mean, can be determined. This probability represents the probability that the individual being considered will choose mode 1.

The last of the 3 analytical techniques is logit analysis. This technique was used by Stopher (26). The modal-choice model that he developed used 2 system characteristics—cost and time—and one user characteristic—income—in the following form:

$$G(x) = \alpha_1(c_1 - c_2) + \alpha_2(t_1 - t_2) + \alpha_3$$

where α_1, α_2, and α_3 are each a function of income, and the c's and t's are the costs and times of travel by the 2 modes considered.

When the function G (X) has been calibrated by determining the values of the α's, the function is substituted in the equation of the logit curve, and p_1 is then the probability that an individual will choose mode 1.

Apart from the models described here, a number of additional models have been developed, many of them as a basis of further experimentation with the techniques. However, the models described here are sufficient to indicate the basic types of models that have been developed by using the disaggregate-behavioral approach.

VALUE OF TIME SAVED BY TRIP PURPOSE

Thomas C. Thomas and Gordon I. Thompson,
Stanford Research Institute

Values of travel-time savings are estimated for personal business trips, social-recreational trips, vacation trips, school trips, and work trips by applying the model previously used by the authors to obtain estimates for commuter work trips. Several route and motorist characteristic variables were studied to determine those that have an effect on the valuation of travel-time savings by motorists. Time saved and income were found to be the most important variables. The benefits of travel-time savings are also shown to differ significantly according to trip purpose. The data base consisted of more than 4,100 usable responses collected in 9 different states from motorists who had a choice between a toll route and a free route. The paper includes tables for estimating the dollar values of travel-time savings as a function of both the motorist's income level and the amount of time saved for each of the 5 trip purposes. In addition, a figure shows the percentage of toll route users as a function of the time-saved and income variables.

•ESTIMATES of the value of travel-time savings as a function of the amount of time saved, income level, and trip purpose are presented in this paper. Such values are required for economic analyses of transportation systems. They convert travel-time savings into equivalent dollar values, which can then be compared with construction costs, maintenance costs, and other real or equivalent cost factors.

The estimated values are the product of a series of studies that were begun in 1962 for the Bureau of Public Roads. The first study developed the conceptual models for a value of travel time saved (1). Small-scale survey work and limited modeling led to a full-scale empirical attempt to estimate a value of time for commuters. This work resulted, in 1968, in an estimate of $2.82 per person per hour as the value of time saved for commuters (2, 3).

However, even as this estimate was made, it was highly qualified. Past work done by Stanford Research Institute (SRI), both theoretical and empirical, indicated that a single constant value of time, even for a single trip purpose, was only a first approximation to a more general variable value. Estimation techniques have subsequently been developed that indicate the value of time saved to be dependent on both the motorist's income level and the amount of time saved. This work was previously reported elsewhere (4).

A second thrust of the study, the principal focus of this paper, was to estimate the value of time saved for trip purposes other than work. Values of time have been estimated for personal business trips, social-recreational trips, vacation trips, and school trips. In addition, the value of time saved for work trips has been revised by use of an enlarged data base.

The purpose of this study, then, has been to determine those route and motorist characteristic variables that have an effect on the valuation of travel-time savings by motorists and to estimate quantitatively the effects of those variables for several trip purposes. Consistent with past findings, the most important route characteristic variable is time saved, and the most important motorist characteristic variable is income.

Sponsored by Task Force on Value of Travel Time.

Therefore, the value of travel time saved depends on both the motorist's income and the amount of time saved. The benefits of travel time savings also differ significantly according to trip purpose. The results by trip purpose support the hypothesized S-shaped benefits curve presented in an earlier report, especially for small time savings (1, 2, 4). Tables for estimating these dollar values of travel time savings as a function of both income and time saved are given in this paper.

The effects of many other variables on the benefits of time saved were studied. Their inclusion in the benefit estimates was not indicated. For example, the effect of an hour-of-day variable appears to be reflected already in the amount-of-time-saved variable, i.e., travel times on alternative routes vary during different hours of the day. The effect of geographic region, i.e., north versus south or urban versus rural, appears to be accounted for by variations reflected in income level so that the estimated benefits can be used nationwide. It was not possible to analyze the effect of other variables, such as type of highway, because of data limitations.

ROUTE-CHOICE MODEL

On the basis of both theoretical and empirical considerations, SRI chose to define the value of savings in travel time in terms of the motorist's perception of costs and time savings. In essence, the perceived costs and other route characteristics are the variables in the route-choice model in the motorist's head. The use of variables not perceived by the motorist has been viewed as a specification error in a route-choice model. This point is emphasized because SRI's premises have led to a different treatment of the measurement of time saved and of changes in motor vehicle operating costs from that used in most other studies (2).

The mathematical model was built by using microeconomic theory. The value of time saved was conceptualized in terms of a motorist's indifference curve for a choice between 2 alternative routes. In microeconomic terminology, the value of time is the slope of the motorist's indifference curve for differential trip costs and trip times; i.e., it is the rate at which the motorist is willing to trade more money for less travel time.

In the real-life situations in which motorists face different route choices, the money and time-saved trade-off cannot be directly observed. It must be inferred from the relationships that emerge when the route choices are estimated on the basis of data on alternative trip costs, time saved, and other route characteristics and on the characteristics of the motorists themselves. The coefficients of the route and the motorist variables in the route-choice estimator specify the relative importance of each variable to the motorist's choice. Therefore, they can be used to calculate the trade-off between cost and time saved.

The mathematical formulation of the route-choice model treated each driver as a separate data point with a binary choice between 2 routes. Monetary considerations in the choice of route were stressed by having 1 route always be a toll road and the other always be a free road.

The analysis estimated the route choice by use of a logit function (2):

$$p(x) = \frac{e^{f(x)}}{1 + e^{f(x)}}$$

where

$p(x)$ = probability of the free road being taken,
$f(x)$ = function of the characteristics of the route and the motorist, and
e = base of the natural logarithms.

The $f(x)$ was restricted to a linear function of the characteristics of the route and the motorist:

$$f(x) = a_0 + a_1x_2 + \ldots + a_nx_n$$

where the a_i's are coefficients to be estimated; and the x_i's are characteristics of the motorist (such as income or sex) or of the route (such as travel time, toll cost, or number of speed-change units).

The p(x) can be interpreted in 2 ways: as a probability estimate of the individual motorist's choice or as the percentage split of a group of motorists all having the same characteristics (the same values) of the independent variables.

When the parameters of the route-choice model have been estimated, the value of travel time saved can easily be estimated from the function f(x). The motorist is assumed to be indifferent to the choice of alternative routes when p(x) equals 0.5. For $p(x) = 0.5$, $f(x) = 0$, or

$$0 = a_0 + a_1x_1 + a_2x_2 + \ldots + a_nx_n$$

This equation can be solved for the sum of monetary charges that the motorist will pay. These charges will then be a function of his characteristics and the characteristics of the alternative routes. The resulting function, called the benefits function, gives the differential amount the motorist will knowingly pay to take the better route. In route-choice situations, it has been demonstrated that motorists neither perceive accurately nor make choices based on the differences in motor vehicle operating costs between alternative routes. Therefore, the only monetary variable in the route-choice equation is the toll charge. If we let x_1 be the toll variable, the amount the motorist is willing to pay is

$$\text{Benefits function} = x_1 = \frac{-a_0}{a_1} - \frac{a_2}{a_1}x_2 - \ldots - \frac{a_n}{a_1}x_n$$

The value of time saved is the derivative of the benefits function with respect to the amount of travel time saved. This derivative is the trading ratio between toll and time saved at a point on the average motorist's indifference curve.

The benefits function finally used in estimating the value of time took 2 general forms:

$$\text{Benefits function} = a_0' + a_1' \Delta t + a_2' I\Delta t$$

and

$$\text{Benefits function} = a_0'' + a_2'' I\Delta t$$

where I is the family income of the driver and Δt is the travel-time difference between alternative routes. This formulation made the value of time saved a function of income level. Estimates were made on data stratified on amount of time saved to establish the dependence of the value of time saved on amount of time saved. The coefficients estimated for each trip purpose are in Appendix B[1]. Appendix C provides estimates from route-choice models incorporating a variety of other route and motorist characteristic variables.

DATA COLLECTION

Personal business, social-recreational, and vacation travel were selected as the most important new trip purposes on which to collect data. Data on school trips and additional data on work trips were also collected.

[1] The original manuscript of this paper included Appendix B, Analyses, Appendix C, Effects of Route and Motorist Characteristic Variables, and Appendix D, Derivations of Average Total Benefits Function. The three appendixes are available in Xerox form at cost of reproduction and handling from the Highway Research Board. When ordering, refer to XS-37, Highway Research Record 369.

SRI investigated the possible advantages of using modal-choice decisions in addition to route-choice decisions to estimate the value of time saved. This option was very attractive because better information on intermodal values of time saved is badly needed. However, it appeared that the quality of the available experimental situations, except for rail rapid transit or railroad, was low. In particular, except for New Orleans, which has special express bus lanes on major streets, it did not appear that nonrail alternatives to the automobile would offer the prerequisite time advantage for a cost and time-saved trade-off analysis.

Therefore, it was decided not to use observations on modal choice and, thus, avoid a sizable number of practical and theoretical questions that had not previously been dealt with. Lisco (5) has been using techniques similar to those used by SRI to study the rail versus automobile alternative.

Sites were selected where motorists on other than work trips would be faced with a choice between a faster toll road and a slower free road. Site locations were found in Florida, Texas, Oklahoma, Maine, New Jersey, Pennsylvania, Virginia, Kentucky, Kansas, and Illinois.

Information was collected by using prepaid mail-back questionnaires designed to require a minimum amount of effort on the respondent's part. The use of such a questionnaire requires English literacy of the driver or a passenger, and this could be a source of bias in the data. However, because the data are stratified by income in the analysis, any bias that might exist is probably limited to those in income level 1, i.e., less than $4,000 per year. Any effect on the overall analysis should be minimal.

Table 1 gives the distribution of motorist and route characteristics by trip purpose. A high proportion of the usable data were for work trips, despite choice of off-peak hours and weekends for most of the survey work. This result was not unexpected, however.

Because the quality of the data obtained compared very favorably with the characteristic of the high-cost personal interview and independent measurement of route characteristics used previously, it was possible to use this low unit cost of data collection confidently to increase the size of the usable data sample by an order of magnitude and explore trip alternatives not previously possible.

BENEFITS OF TIME SAVED

The benefits of time saved are given in Table 2 for work trips, in Table 3 for social-recreational trips, in Table 4 for personal business trips, in Table 5 for vacation trips, and in Table 6 for school trips.

The income variable is the family income of the driver. For social-recreational, personal business, and vacation trips where those in the car are usually members of the same family, the best estimates of benefits were obtained on a per-vehicle basis. For work and school trips, the benefits were estimated on a per-person basis, i.e., per vehicle-passenger. School trips were estimated almost entirely for college students.

The tables give the total value of the specific amount of time saved rather than the value per hour. These dollar figures for specific amounts of time are called the benefits of time saved, whereas dollars-per-hour figures are called the values of time saved. For example, the benefits of a 15-min time saving during work trips to motorists earning $8,000 to $10,000 per year are valued at 57.0 cents (Table 2); i.e., the average amount motorists would be willing to pay for a 15-min time saving is 57.0 cents. (Note that this amount is not the hourly value of time saved nor the value of the 15th minute by itself but rather the sum of the values of each of the 15 individual minutes.)

The value of time saved is a nonlinear function of the amount of time saved, so benefits are not a linear function of the amount of time saved. Hence, the benefits of an amount of time saved cannot be obtained by simply multiplying the amount of the time saving by a constant value of time saved (e.g., the $2.82 per person per hour previously estimated for commuters).

TABLE 1

NUMBER OF USABLE RESPONSES OR DATA POINT STRATIFICATION BY TRIP PURPOSE

	Toll Route Data					Free Route Data				
	Work	School	Vacation	Personal Business	Social-Recreational	Work	School	Vacation	Personal Business	Social-Recreational
Toll cost (cents)										
0-10	47	3	3	14	8	25	2	0	18	20
11-20	152	3	7	33	31	36	3	3	25	18
21-30	172	5	24	48	89	94	12	8	71	100
31-40	184	8	10	47	42	104	7	4	37	42
41-50	174	5	9	41	43	71	17	16	31	39
51-60	71	12	21	38	32	37	18	2	33	29
61-70	125	8	13	39	43	41	2	2	20	12
71-80	60	3	3	21	25	20	11	4	19	26
81-90	30	5	5	13	13	16	6	5	8	11
91-100	42	2	15	18	34	18	2	24	6	21
101-150	197	8	44	63	105	54	5	35	39	54
151-200	41	2	21	21	39	8	2	5	10	12
201-250	29	0	21	22	21	4	3	9	7	10
Over 250	48	5	44	35	29	4	0	10	7	6
Time saved (min)										
0-5	80	0	3	19	17	132	13	3	47	53
6-10	310	20	17	73	74	192	38	13	122	101
11-15	314	14	19	85	121	96	27	25	77	105
16-20	219	7	25	59	65	44	8	21	29	56
21-25	44	5	5	17	25	9	2	6	7	11
26-30	213	10	50	77	121	40	1	29	25	45
31-35	17	0	3	5	4	2	0	3	5	1
36-40	25	0	6	12	16	5	1	2	4	4
41-45	43	1	15	27	21	5	0	4	5	10
46-50	5	1	3	3	5	0	0	0	1	1
51-55	1	1	1	4	0	0	0	0	0	0
56-60	65	5	48	45	46	4	0	18	7	11
61-90	14	2	17	13	12	1	0	1	1	1
91-120	19	2	15	9	20	2	0	2	1	1
Over 120	3	1	13	5	7	0	0	0	0	0
Trip frequency										
Daily	583	24	2	27	15	293	58	1	18	15
Once a week or more	343	24	22	109	101	111	21	20	86	115
Once a month or more	192	3	18	97	114	59	5	15	73	81
Occasionally	203	14	144	186	270	52	5	68	122	155
First time	51	4	54	34	54	17	1	23	32	34
Time of day										
7:00-7:59	157	2	0	6	0	23	1	1	3	1
8:00-8:59	197	26	15	36	17	97	33	3	31	23
9:00-9:59	86	7	24	42	55	43	3	4	24	24
10:00-10:59	174	6	26	77	63	95	11	23	52	65
11:00-11:59	17	1	9	10	23	16	7	0	8	2
12:00-12:59	65	4	17	23	29	27	2	3	26	36
1:00-1:59	145	1	25	58	65	68	3	27	56	69
2:00-2:59	151	9	34	60	51	27	15	15	29	34
3:30-3:59	129	3	20	37	38	49	4	5	38	41
4:00-4:59	163	4	16	40	59	51	9	24	34	52
5:00-5:00	15	3	14	11	31	4	0	8	5	15
6:00-6:59	11	1	20	10	40	6	0	8	8	23
7:00-7:59	47	2	14	32	70	25	2	2	15	11
8:00 and later	15	0	6	11	13	1	0	4	2	4
Day of week										
Monday	185	6	12	54	22	49	7	2	23	14
Tuesday	192	4	10	52	23	55	7	3	22	12
Wednesday	343	17	8	39	31	101	10	2	29	18
Thursday	190	9	13	40	14	121	39	6	44	25
Friday	205	13	49	80	109	125	20	34	83	74
Saturday	225	14	92	133	193	68	6	55	107	157
Sunday	32	6	56	55	162	13	1	25	23	100

TABLE 1 (Concluded)

	Toll Route Data					Free Route Data				
	Work	School	Vacation	Personal Business	Social-Recreational	Work	School	Vacation	Personal Business	Social-Recreational
Route reason										
Less congestion	97	0	9	35	35	18	0	13	15	17
Shorter time	542	23	89	153	185	0	0	0	0	0
Safety	38	4	12	20	21	4	0	6	8	2
No toll	0	0	0	0	0	298	68	42	158	190
Lower gas and oil costs	3	1	0	1	1	2	0	0	1	3
Scenery	3	0	3	1	7	19	2	17	17	29
Convenience	121	6	32	43	51	119	4	19	72	80
Availability of disabled vehicle service	0	0	1	1	1	7	0	2	8	3
Combinations of all these reasons	568	35	94	199	253	65	74	90	279	324
Number of adults										
1	1,147	45	40	223	158	449	66	21	164	127
2	182	17	155	190	305	63	18	81	134	203
3	27	3	28	28	44	14	3	14	22	41
4	8	0	14	7	31	4	1	8	10	23
5	3	4	3	4	10	1	1	2	0	5
6	3	0	0	1	5	0	0	1	1	1
7	2	0	0	0	1	1	1	0	0	0
Automobile year										
1950 or earlier	1	0	0	0	0	1	0	0	1	1
1951-55	2	1	0	3	2	1	2	1	2	1
1956-60	24	1	2	8	14	5	7	0	10	17
1961-65	232	21	67	101	146	141	24	48	113	126
1966-70	1,113	46	171	341	392	384	57	83	205	255
Income level (per year)										
Under $4,000	15	9	10	18	30	11	15	5	29	46
$4,000-5,999	37	2	16	36	36	23	9	9	43	41
$6,000-7,999	92	8	19	48	60	64	14	12	52	65
$8,000-9,999	156	8	29	42	61	84	19	27	38	49
$10,000-11,999	198	7	36	62	87	87	8	15	45	67
$12,000-14,999	299	11	50	70	81	113	10	15	60	54
$15,000-19,999	244	10	32	72	83	76	7	12	27	43
$20,000 or more	331	14	48	105	116	74	8	32	37	35
Driver's age										
Under 19	7	9	2	5	7	5	6	4	7	16
19-25	116	20	29	63	101	65	50	18	68	93
26-65	1,232	40	202	359	417	457	34	101	238	267
Over 65	17	0	7	26	29	5	0	4	18	24
Driver's sex										
Male	1,216	48	210	339	412	468	70	117	230	284
Female	156	21	30	114	142	64	20	10	101	116
Total	1,372	69	240	453	554	532	90	127	331	400

For an example of this nonlinearity in the value of time, consider again the work-trip data given in Table 2. It is estimated that a motorist earning $8,000 to $10,000 per year would be willing to pay 1.9 cents for a 5-min saving, 27.6 cents for a 10-min saving, and 73.2 cents for a 20-min saving. Note that the value of the 20-min saving is not twice that of the 10-min saving, nor is the value of the 5-min saving half that of the 10-min saving.

The tables have been developed subject to 2 limitations—an upper limit on the amount of time saved and, within this limit, a further limitation on the maximum amount of benefits per person or per vehicle. Both of these limitations result from constraints in the range of variables in the data sample and are not theoretically required. The upper limit on time saved is 40 min on work trips, 30 min on personal business and vacation trips, and 20 min on social-recreational and school trips. Within this time restriction, the maximum benefit that should be considered for any single person or

TABLE 2

BENEFITS OF TIME SAVINGS IN DOLLARS PER PERSON FOR WORK TRIPS

Time Saving (min)	Income Level of Motorist							
	1	2	3	4	5	6	7	8
1	0.001	0.002	0.003	0.004	0.006	0.008	0.011	0.015
2	0.002	0.004	0.005	0.008	0.011	0.016	0.022	0.031
3	0.004	0.005	0.008	0.012	0.017	0.024	0.033	0.046
4	0.005	0.007	0.011	0.015	0.022	0.032	0.045	0.061
5	0.006	0.009	0.013	0.019	0.028	0.040	0.056	0.077
6	0.009	0.014	0.022	0.034	0.051	0.075	0.108	0.149
7	0.013	0.022	0.036	0.057	0.089	0.132	0.186	0.249
8	0.018	0.033	0.056	0.093	0.144	0.210	0.285	0.365
9	0.026	0.048	0.086	0.142	0.216	0.302	0.393	0.487
10	0.036	0.070	0.126	0.205	0.299	0.401	0.505	0.610
11	0.050	0.099	0.177	0.276	0.387	0.502	0.618	0.732
12	0.068	0.136	0.236	0.354	0.478	0.604	0.729	0.854
13	0.091	0.180	0.300	0.434	0.570	0.706	0.841	0.975
14	0.119	0.230	0.368	0.514	0.661	0.807	0.952	1.096
15	0.135	0.261	0.412	0.570	0.727	0.883	1.033	1.191
16	0.139	0.273	0.433	0.602	0.768	0.934	1.094	1.261
17	0.144	0.285	0.454	0.633	0.810	0.986	1.155	1.330
18	0.148	0.297	0.476	0.666	0.853	1.038	1.216	1.400
19	0.153	0.309	0.498	0.699	0.896	1.091	1.277	1.470
20	0.158	0.322	0.521	0.732	0.939	1.143	1.338	1.539
21	0.162	0.335	0.544	0.765	0.982	1.195	1.399	1.608
22	0.167	0.348	0.567	0.799	1.025	1.247	1.460	1.678
23	0.172	0.361	0.591	0.833	1.069	1.299	1.521	1.747
24	0.177	0.375	0.615	0.867	1.112	1.352	1.581	1.816
25	0.182	0.389	0.639	0.901	1.156	1.404	1.642	1.885
26	0.187	0.403	0.664	0.935	1.199	1.456	1.702	1.954
27	0.192	0.417	0.688	0.970	1.243	1.508	1.763	2.023
28	0.198	0.431	0.713	1.004	1.286	1.560	1.823	2.091
29	0.203	0.446	0.738	1.039	1.329	1.612	1.884	2.160
30	0.208	0.460	0.763	1.074	1.373	1.664	1.944	2.229
31	0.214	0.475	0.789	1.108	1.416	1.715	2.004	2.298
32	0.219	0.491	0.814	1.143	1.459	1.767	2.064	2.366
33	0.225	0.506	0.839	1.178	1.503	1.819	2.124	2.435
34	0.230	0.521	0.865	1.212	1.546	1.870	2.184	2.504
35	0.236	0.537	0.891	1.247	1.589	1.922	2.244	2.572
36	0.242	0.552	0.916	1.282	1.632	1.974	2.305	2.641
37	0.247	0.568	0.942	1.316	1.675	2.025	2.365	2.709
38	0.253	0.584	0.968	1.351	1.719	2.077	2.425	2.778
39	0.259	0.600	0.994	1.386	1.762	2.128	2.485	2.847
40	0.265	0.616	1.020	1.420	1.805	2.180	2.545	2.915

Annual income level: 1 = less than $4,000; 2 = $4,000-5,999; 3 = $6,000-7,999; 4 = $8,000-9,999; 5 = $10,000-11,999; 6 = $12,000-14,999; 7 = $15,000-19,999; and 8 = more than $20,000.

vehicle is $1 for work trips and $2 for all other trips. Larger dollar amounts are given in the tables ($1 and more than $1, Table 2; $2 and more than $2, Tables 3-5), but the use of these data should be avoided if possible because they are an extrapolation outside the data base.

Limitations placed on the use of the tables and the choice of the specific estimators used to construct them were to a large extent based on the authors' best subjective judgment, given the data available. In most instances, however, either the choices were quite obvious or the results were relatively insensitive to the particular choice.

The dollar amounts in the benefit tables are the total benefits as perceived by the motorist for the faster route. Other benefits calculated by highway engineers such as changes in operating, maintenance, and accident costs are in general not perceived by the nonbusiness motorist. It is recommended that the different types of benefits not be added together in making benefit-cost calculations but rather be shown separately to the highway decision-maker.

TABLE 3

BENEFITS OF TIME SAVINGS IN DOLLARS PER VEHICLE FOR SOCIAL-RECREATIONAL TRIPS

Time Saving (min)	Income Level of Motorist							
	1	2	3	4	5	6	7	8
1	0.000	0.000	0.000	0.001	0.001	0.003	0.005	0.010
2	0.000	0.000	0.001	0.001	0.002	0.005	0.010	0.020
3	0.000	0.000	0.001	0.002	0.004	0.008	0.015	0.030
4	0.000	0.001	0.001	0.002	0.005	0.010	0.020	0.040
5	0.000	0.001	0.001	0.003	0.006	0.013	0.026	0.050
6	0.000	0.001	0.002	0.005	0.013	0.029	0.065	0.132
7	0.000	0.001	0.003	0.009	0.026	0.065	0.146	0.278
8	0.001	0.002	0.005	0.017	0.050	0.132	0.278	0.467
9	0.001	0.002	0.008	0.029	0.094	0.236	0.442	0.669
10	0.001	0.003	0.013	0.050	0.162	0.369	0.618	0.872
11	0.001	0.004	0.019	0.083	0.256	0.517	0.796	1.074
12	0.001	0.005	0.029	0.132	0.369	0.669	0.973	1.274
13	0.001	0.007	0.044	0.197	0.492	0.822	1.149	1.475
14	0.001	0.009	0.065	0.278	0.618	0.973	1.324	1.674
15	0.002	0.014	0.086	0.343	0.719	1.098	1.469	1.839
16	0.003	0.021	0.108	0.392	0.795	1.197	1.585	1.971
17	0.004	0.028	0.131	0.443	0.874	1.296	1.701	2.102
18	0.005	0.035	0.157	0.496	0.954	1.396	1.816	2.233
19	0.006	0.044	0.184	0.552	1.035	1.495	1.932	2.364
20	0.007	0.052	0.213	0.611	1.117	1.595	2.047	2.495

TABLE 4

BENEFITS OF TIME SAVINGS IN DOLLARS PER VEHICLE FOR PERSONAL BUSINESS TRIPS

Time Saving (min)	Income Level of Motorist							
	1	2	3	4	5	6	7	8
1	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003
2	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.007
3	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.010
4	0.000	0.000	0.000	0.000	0.001	0.002	0.005	0.014
5	0.000	0.000	0.000	0.000	0.001	0.002	0.006	0.017
6	0.000	0.000	0.000	0.001	0.002	0.007	0.026	0.084
7	0.000	0.000	0.000	0.001	0.006	0.026	0.100	0.272
8	0.000	0.000	0.001	0.003	0.017	0.084	0.272	0.543
9	0.000	0.000	0.001	0.007	0.048	0.213	0.508	0.827
10	0.000	0.000	0.002	0.017	0.118	0.403	0.756	1.108
11	0.000	0.000	0.004	0.039	0.241	0.614	1.003	1.388
12	0.000	0.001	0.007	0.084	0.403	0.827	1.248	1.667
13	0.000	0.001	0.014	0.161	0.579	1.038	1.492	1.946
14	0.000	0.001	0.026	0.272	0.756	1.248	1.737	2.225
15	0.001	0.005	0.058	0.402	0.950	1.478	2.004	2.480
16	0.002	0.012	0.109	0.552	1.160	1.728	2.293	2.811
17	0.002	0.020	0.175	0.713	1.370	1.977	2.583	3.141
18	0.003	0.032	0.258	0.879	1.579	2.225	2.872	3.472
19	0.003	0.046	0.353	1.047	1.787	2.473	3.161	3.802
20	0.004	0.065	0.463	1.215	1.994	2.721	3.450	4.132
21	0.005	0.089	0.580	1.383	2.201	2.969	3.739	4.463
22	0.007	0.119	0.702	1.550	2.408	3.217	4.028	4.793
23	0.008	0.155	0.827	1.716	2.614	3.465	4.318	5.124
24	0.009	0.198	0.952	1.882	2.821	3.712	4.607	5.454
25	0.011	0.249	1.079	2.048	3.027	3.960	4.896	5.784
26	0.013	0.306	1.205	2.214	3.234	4.208	5.185	6.115
27	0.016	0.370	1.330	2.379	3.440	4.456	5.474	6.445
28	0.018	0.439	1.456	2.544	3.647	4.704	5.763	6.776
29	0.022	0.513	1.581	2.710	3.854	4.951	6.052	7.106
30	0.025	0.590	1.706	2.875	4.060	5.199	6.341	7.437

TABLE 5

BENEFITS OF TIME SAVINGS IN DOLLARS PER VEHICLE FOR VACATION TRIPS

Time Saving (min)	Income Level of Motorist							
	1	2	3	4	5	6	7	8
1	0.044	0.050	0.057	0.064	0.073	0.082	0.091	0.102
2	0.087	0.100	0.113	0.129	0.145	0.163	0.183	0.203
3	0.131	0.149	0.170	0.193	0.218	0.245	0.274	0.305
4	0.174	0.199	0.227	0.257	0.291	0.327	0.365	0.407
5	0.218	0.249	0.284	0.322	0.363	0.408	0.457	0.509
6	0.224	0.262	0.306	0.355	0.408	0.467	0.530	0.598
7	0.230	0.276	0.330	0.390	0.457	0.530	0.609	0.694
8	0.236	0.291	0.355	0.427	0.509	0.598	0.694	0.796
9	0.242	0.306	0.381	0.467	0.563	0.669	0.783	0.903
10	0.249	0.322	0.408	0.509	0.621	0.744	0.875	1.013
11	0.256	0.338	0.437	0.552	0.681	0.822	0.971	1.126
12	0.262	0.355	0.467	0.598	0.744	0.903	1.069	1.241
13	0.269	0.372	0.498	0.645	0.809	0.985	1.169	1.357
14	0.276	0.390	0.530	0.694	0.875	1.069	1.270	1.474
15	0.283	0.408	0.562	0.741	0.940	1.152	1.370	1.590
16	0.290	0.425	0.593	0.788	1.004	1.233	1.468	1.705
17	0.297	0.443	0.624	0.835	1.069	1.315	1.568	1.821
18	0.305	0.461	0.656	0.884	1.135	1.399	1.669	1.938
19	0.312	0.480	0.689	0.934	1.202	1.484	1.770	2.055
20	0.320	0.499	0.723	0.985	1.271	1.569	1.872	2.172
21	0.327	0.518	0.758	1.036	1.340	1.656	1.974	2.290
22	0.335	0.538	0.793	1.089	1.410	1.742	2.076	2.407
23	0.343	0.558	0.829	1.143	1.481	1.830	2.179	2.524
24	0.350	0.579	0.866	1.197	1.552	1.917	2.282	2.641
25	0.358	0.599	0.903	1.252	1.624	2.005	2.384	2.758
26	0.367	0.621	0.941	1.307	1.696	2.093	2.487	2.875
27	0.375	0.642	0.979	1.363	1.769	2.181	2.589	2.992
28	0.383	0.644	1.018	1.419	1.842	2.269	2.692	3.109
29	0.392	0.687	1.058	1.576	1.915	2.357	2.794	3.225
30	0.400	0.709	1.089	1.533	1.988	2.445	2.897	3.341

TABLE 6

BENEFITS OF TIME SAVINGS IN DOLLARS PER PERSON FOR SCHOOL TRIPS

Time Saving (min)	Income Level of Motorist							
	1	2	3	4	5	6	7	8
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
2	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001
3	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002
4	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.003
5	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.004
6	0.000	0.000	0.000	0.000	0.001	0.002	0.005	0.011
7	0.000	0.000	0.000	0.001	0.002	0.005	0.013	0.032
8	0.000	0.000	0.000	0.001	0.004	0.011	0.032	0.083
9	0.000	0.000	0.001	0.002	0.007	0.025	0.075	0.179
10	0.000	0.000	0.001	0.004	0.014	0.053	0.151	0.309
11	0.000	0.000	0.001	0.006	0.028	0.013	0.258	0.452
12	0.000	0.000	0.002	0.011	0.053	0.179	0.380	0.597
13	0.000	0.000	0.003	0.019	0.093	0.275	0.507	0.741
14	0.000	0.001	0.005	0.032	0.151	0.380	0.633	0.884
15	0.000	0.001	0.007	0.053	0.225	0.489	0.759	1.027
16	0.000	0.001	0.011	0.083	0.309	0.597	0.884	1.169
17	0.000	0.002	0.017	0.126	0.398	0.705	1.009	1.312
18	0.000	0.002	0.025	0.179	0.489	0.813	1.134	1.454
19	0.000	0.003	0.036	0.241	0.579	0.920	1.258	1.596
20	0.000	0.004	0.053	0.309	0.669	1.027	1.383	1.739

As an added aid to the highway economist, graphs and equations have been prepared that indicate the percentage of toll route users as a function of the amount of time saved and the toll costs (Appendix).

A CHALLENGE

The analyst who attempts to use these tables will find a whole new set of requirements for data on highway improvements as a result, primarily, of the finding that the value of time is a function of the amount of time saved. The total amount of time saved by a motorist on his trip is therefore crucial. The average value of 1 min of time saved is dependent on whether this minute is the only time saved, part of a total of 10 min saved, or part of a total of 20 min saved. As an example, at income level 4 (Table 2), 1 min will be valued at 0.4 cents if it is the only minute saved, at an average of 2.05 cents (20.5/10 = 2.05) if it is part of a 10-min saving, or at an average of 3.66 cents (73.2/20 = 3.66) if it is part of a 20-min saving.

The highway economist now requires more information than just the amount of time saved by a single highway improvement and the volume of motorists who will use it. He needs cross-tabulated information on all the improvements (actual and planned), on motorists' different trips and their trip lengths, and on their income levels.

Consequently, the use of these tables imposes data requirements far greater than those met by data currently collected for highway economy studies. Yet even the present requirements, such as estimates for the volume of motorists who will use a single improvement, are sometimes difficult to meet accurately. The result is certain to present the highway economist with a challenge.

REFERENCES

1. Haney, D. G. The Value of Time for Passenger Cars, Volume I: A Theoretical Analysis and Description of Preliminary Experiments. Stanford Research Institute, Menlo Park, Calif., May 1967.
2. Thomas, T. C. The Value of Time for Passenger Cars, Volume II: An Experimental Study of Commuter's Values. Stanford Research Institute, Menlo Park, Calif., May 1967.
3. Thomas, T. C. Value of Time for Commuting Motorists. Highway Research Record 245, 1968, pp. 17-35.
4. Thomas, T. C., and Thompson, G. I. The Value of Time for Commuting Motorists as a Function of Their Income Level and Amount of Time Saved. Highway Research Record 314, 1970, pp. 1-19.
5. Lisco, T. E. The Value of Commuters' Travel Time—A Study in Urban Transportation. Univ. of Chicago, 1967; abridged in Highway Research Record 245, 1968, p. 36.

APPENDIX

TOLL-ROUTE PATRONAGE ESTIMATES

The percentage of toll-route users P can be determined from

$$P = [1 + \exp(F)]^{-1}$$

where F is the estimated discriminant function for the trip purpose of interest. The discriminant functions to be used in this equation are given in Table 7 for each trip purpose. (Although these estimated functions have been modified for fitting the benefits functions more closely to the data, the constant terms have not been adjusted to eliminate the discontinuity encountered at the point of time savings where the 2 piecewise functions meet.)

Figure 1 shows the percentage of toll-route users as a function of both time saved and toll cost for work trips. Note that demand is inelastic for the higher income levels and larger time savings. Smaller time savings and lower income levels yield more elastic demand curves. Similar curves and results can be obtained for the other trip purposes.

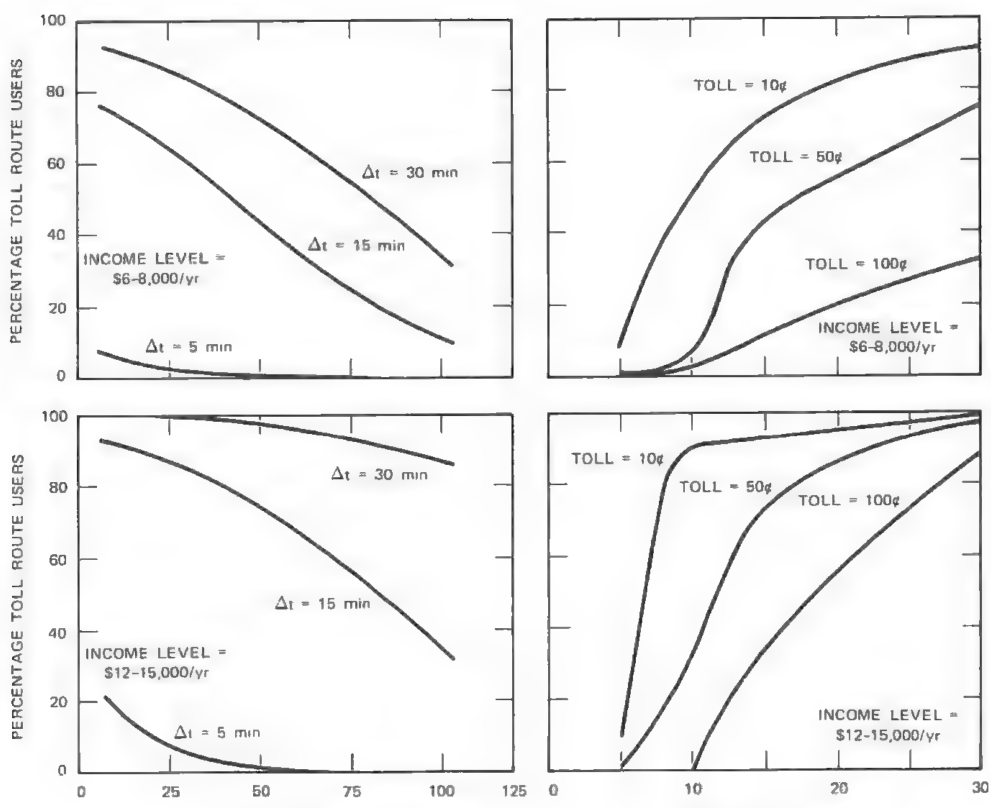


Figure 1. Toll-route patronage for work trips

TABLE 7

ESTIMATED DISCRIMINANT FUNCTIONS BY TRIP PURPOSE

Trip Purpose	Discriminant Function F	Time Interval (min)
Work	$+4.49 - 0.290\,\Delta t - 0.778\,I\Delta t + 0.0757\,C$	5-15
	$-0.171 + 0.0014\,\Delta t - 0.0290\,I\Delta t + 0.0337\,C$	5-30
Personal business	$+8.50 - 0.158\,I\Delta t + 0.0627\,C$	5-15
	$+3.16 - 0.0975\,I\Delta t + 0.03660\,C$	5-30
Social-recreational	$+6.66 - 0.148\,I\Delta t + 0.0594\,C$	5-15
	$+2.08 - 0.0457\,I\Delta t + 0.0281\,C$	5-20
School	$+13.1 - 0.188\,I\Delta t + 0.0813\,C$	5-20
Vacation	$+0.275 - 0.0332\,I\Delta t + 0.231\,C$	5-30

DISCUSSION

Shalom Reichman, Hebrew University, Jerusalem, Israel

It is clear that some of the contributions of this paper are likely to have noticeable effects in the field of transportation economics. In terms of modeling transportation demand, for instance, the use of perceived data alone to generate time values, the construct of time benefits as distinct from time values, and the derivation of different time benefits by trip purposes will each have important implications. For this reason, it is appropriate to raise a number of questions concerning the applicability of the reported results.

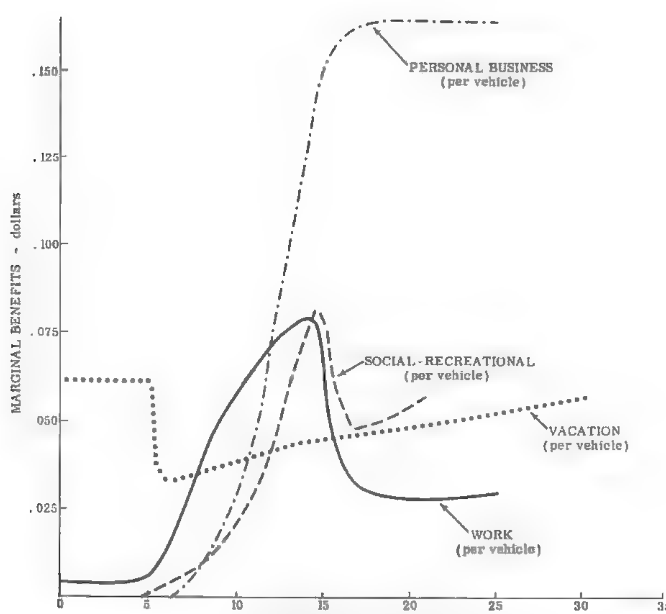


Figure 2. Marginal benefits of time saved, by purpose and amount of time saved, for annual income level of $8,000-9,999.

One question refers to an apparently unrealistic use of perceived data. Let us assume, as did the authors, that a behavioral microeconomic approach to route choice requires the use of perceived rather than measured data. But then to assume a continuous, minute-by-minute perception of time saved is not realistic because such a brief time interval reaches the limit of human discrimination. By plotting the marginal benefits of time saved by income class ($\Delta b/\Delta t$, $Y = Y_0$) as given in Tables 2, 3, 4 and 5, one can compare the results in a diagrammatic form (Fig. 2). It should be noted that Figure 2 shows only those values within the original constraints of 20, 30, and 40 min of time saved for social-recreational, personal business and vacation, and work trips respectively. The same holds for the maximum benefit restriction of $1 for work trips and $2 for all other trips. Also, school trips were not included, as the authors themselves place a wide confidence limit on the derived values.

It is immediately apparent that the time-saved continuum, or x-axis, can be divided into 3 zones. In the first zone, ranging from 0 to 5 min saved, marginal time benefits for each purpose are constant, although of different magnitudes. The same is generally true for time savings of more than 16 min. Only in the second zone, which ranges from 5 to 16 min, do marginal benefits change as a function of the amount of time saved. A question that naturally presents itself is whether this segregation of the time continuum stems from empirical observations in the field or from mathematical requirements of the models used in the derivation of time benefits. Moreover, no behavioral or economic explanation is offered for the different shapes and magnitudes of the marginal benefits, even in the interval of 5 to 16 min saved.

To conclude, it is suggested that a general application of the reported results should await further clarifications on the points raised in the preceding.

AUTHORS' CLOSURE

The discussion has raised some technical questions about the research. It is clear that other questions could also be raised and that future research will improve on the results presented in this paper. However, the crucial question for the highway economist is, should the present results be used in making economic analyses? Reichman's answer appears to be that they should not be used without further clarification of some technical points.

However, if these results are not used, then the "old results" (also developed by SRI for the Federal Highway Administration) are all that are available (2). As Reichman knows, it was the deficiencies in these old results that lead to the present study and the results reported here. The deficiencies of the previous work included the limitation of trip purpose to commuters, a linearization of the value of time saved that was not supported by the data, a value of time not directly a function of income level, and a much smaller and more geographically limited data base. These deficiencies were overcome in the new study. There is no question in our minds that, subject to the limitations set forth in the paper, the new results represent the most accurate inputs to highway economy studies currently available.

The figures that Reichman plotted appear to be consistent with the S-shaped benefits function that is supported by both theory and empirical evidence (2, 3, 4). In general, piecewise linear estimates were made on 2 intervals—5 to 15 min and greater than 15 min. Continuity requirements were used to estimate the value of time saved in the 0-to-5-min interval. The precise estimates for each trip purpose are available in Appendix B. The development of the techniques is given in a previous publication (4).

A second point concerns the use of motorist-reported time savings. The issue of reported versus measured versus perceived data (which, it should be noted, are not the same as reported data) has concerned us throughout our series of studies (2). It is a complicated question involving matters of accuracy of each sample point, sample size, and geographic coverage within the constraints of limited financial resources. Our analyses of previous data samples involving both measured and reported data on the

same trips indicated that reported data were at least as valid as measured data (4). It is incorrect of Reichman to state that an unrealistic continuous minute-by-minute perception of time saved was made. Neither we nor the mathematical techniques employed make such an assumption.

We hope that the paper will raise many methodological and empirical questions in the field of transportation economics. However, because practice cannot wait until all research questions are answered, we recommend that the reported results be put to use now by practitioners.

DEVELOPMENT OF A DOWNTOWN PARKING MODEL

Gökmen Ergün*, Turkish Highway Department, Ankara

Provision of adequate downtown parking is a serious problem facing most urban areas. In this study, probabilistic models have been built to explain the behavior of people in choosing among parking places. The effects on parking choices of parking system characteristics and socioeconomic factors have been investigated. The analytical tool used was logit analysis. In the study, implications were drawn and an example worked out for an application of the models. Methods were devised for finding the value of time saved walking from parking place to destination. The methods were applied to data, and values of time were established that are suitable for cost-benefit analysis of downtown parking facility investments. The behavioral models developed can be used to investigate expected parking costs and parking cost changes, and it is possible to assign persons to different parking distances from destinations.

THE ANALYSIS of downtown parking behavior investigates what determines where people park. The basic trade-offs involved are essentially those of choosing between parking closer to destinations at higher money cost and lower walking time and parking farther away at lower cost but with a longer walk. From the analysis of actual parking decisions made, it is possible both to predict what persons would do given alternative future situations and to infer traveler values associated with choosing between walking farther and paying more. These predictions and values can then be used in planning and evaluating parking projects and policies.

Some of the particular questions that the analysis of downtown parking behavior can be helpful in answering are as follows:

1. What are the effects of changes in parking price policy? As prices change, how do drivers correspondingly change where they park?
2. Given the costs, what are the benefits of investments in new parking facilities? Where should parking be located to best serve needs? What kind of parking facilities should be provided? What should parking fees be?

In this study, an analysis of downtown parking behavior has been made in which a number of probabilistic models were built to help explain traveler choices among parking places. Explanatory variables investigated included system characteristics (e.g., parking cost and parking distance), socioeconomic factors (e.g., income, age, and sex), and trip purpose (work and nonwork). The analytical tool used was logit analysis.

In the study, implications have been drawn and some examples worked out for applications of the models. Methods were devised for finding the value of time saved walking from parking place to destination. The methods were applied to data, and values of time were established that are suitable for cost-benefit analysis of downtown parking facility investments. The behavioral models developed can be used to find expected parking costs and parking times to investigate the effects of parking cost changes, and to assign persons to different parking distances from destinations.

For development of the models, data collected for the Northwest Chicago Corridor Modal-Split Project were used (6). These statistics include data on 226 downtown

*Mr. Ergün was with the Chicago Area Transportation Study when this research was performed.

Sponsored by Task Force on Value of Travel Time.

automobile trips and information on socioeconomic factors, alternative parking costs, downtown destinations, and trip purposes. The trips from which these data were derived originated in 2 areas of the northwest Chicago travel corridor. The first area is suburban Park Ridge and a small portion of northwest Chicago. The second combines the contiguous suburbs of Arlington Heights and Rolling Meadows.

Of the total sample of 226 downtown automobile trips, 41 involved company-paid drivers, 7 involved riders, and 13 had unreasonable or miscoded information.

For analysis purposes, data for only 165 individuals were used after data involving company-paid drivers, riders, and unreasonable or miscoded samples were eliminated. Of the net sample of 165 individuals making trips, 113 involved work trips, and 52 involved nonwork trips. The data analyzed included both persons who drove by choice (i.e., who could have taken the train) and those who needed to drive.

STATISTICAL METHOD

Although both probit (2, 5, 9) and logit (8, 10) methods were suitable for the statistical analysis, logit analysis was used. To understand the logit method as it relates to the parking problem, assume that the objective is to build a model that finds the probability, p, that a driver will choose to park at a given distance from his destination. Assume also that $X_1, X_2, \ldots X_n$ are the variables influencing the choice. Then, in a very simple form, it can be hypothesized that

$$p = a_0 + a_1 X_1 + a_2 X_2 + \ldots a_n X_n \tag{1}$$

But this form is inappropriate because it can take values between $-\infty$ and $+\infty$, whereas probability p is constrained to values between 0 and 1. This can be avoided by using another functional form:

$$\frac{p}{1 - p} = e^{a_0} X_1^{a_1} X_2^{a_2} \ldots \tag{2}$$

Note that as $p \to 0$, and as $1 - p \to 1$, $p/(1 - p) \to \infty$. Taking the logs of both sides, we get

$$\mathrm{Log}\left(\frac{p}{1 - p}\right) = a_0 + a_1 \log X_1 + a_2 \log X_2 \ldots \tag{3}$$

Also note that for the different values of $p/(1 - p)$, $\log p/(1 - p)$ varies between $-\infty$ and $+\infty$ and is a monotonically increasing function of the probability p. The expression $\log p/(1 - p)$ is known as the logit of the probability p.

If the right side of Eq. 2 is replaced by a function G(X), it becomes

$$\mathrm{Log}\left(\frac{p}{1 - p}\right) = G(X) \tag{4}$$

hence,

$$\frac{p}{1 - p} = e^{G(X)} \tag{5}$$

and

$$p = \frac{e^{G(X)}}{1 + e^{G(X)}} \tag{6}$$

This expression yields a sigmoid (or S) curve that is symmetrical and is similar to the cumulative normal curve, diverging from it at the extremes only (Fig. 1). The coefficients of Eq. 3 can be found by maximum likelihood estimation (8).

PROBLEM FORMULATION

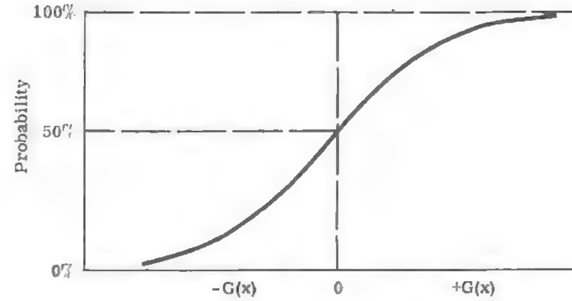


Figure 1. Sigmoid curve.

The basic analytical assumption used in this study was that a person will always choose the parking location that best satisfies his needs and desires. He evaluates parking locations on the basis of parking fee, distance of the parking from his destination, and slope of the parking cost profile. Trip purpose is also relevant.

Based on the preceding assumption, the approach used was to consider parking separately for each block distance and to find the probabilities of using such parking for each individual. When there were multiple parking places within a block, the cheapest one was considered. In other words, a binary-choice model was built for each parking distance, the choices being to use or not to use parking at the distance under consideration. This approach may be a good simulation of the real choice procedure. It seems reasonable to assume that an individual first considers the nearest parking location, and, if it is convenient from his point of view and compares favorably with other choices, he will use it. If not, he will go to the next nearest location and value it. This process will continue until he finds the location that has the greatest value to him.

In the analysis of choices made, 2 types of possible explanatory variables were investigated. The first type that represented system characteristics included parking cost, parking distance, slope of the parking cost profile, and number of parking hours. The second type of explanatory variable investigated was socioeconomic characteristics of the individual choosing parking. Characteristics analyzed were income, age, and sex. Analysis was conducted separately for work and nonwork trips.

Each of the variables considered in the analysis and the way individual variables were handled are discussed in the following sections.

Parking Cost

Parking cost is perhaps the most obvious variable to investigate. As the parking cost increases at a certain location, one would expect that it becomes less likely for persons to park there. Presumably, if the parking cost is high enough, people will prefer to walk a longer distance rather than to pay a higher price to park.

Parking Distance

The distance variable may be expected to have a negative effect on the choice in the sense that, as distance from a parking location to destination increases, it becomes less likely for people to park there. This variable was taken into account by building a separate logit model for each distance interval. Through so doing, the effect of distance was reflected in the constant terms of the logit functions.

Slope of Parking Cost Profile

The variable of the cost-profile slope represents the effect of parking cost differences on people's choices. To see this more clearly, consider the 2 parking cost profiles given as follows:

	Block				
Profile	First	Second	Third	Fourth	Fifth
1	\$2.50	\$2.45	\$2.40	\$2.40	\$2.35
2	\$3.75	\$3.00	\$2.00	\$1.00	\$0.25

In the first profile, the parker must walk 5 blocks to save 15 cents. In the second, if he walks 5 blocks, he will save \$3.50. Naturally, if the saving is small compared to the parking cost (as in the first example) and is also small with respect to the individual's value of time, it is reasonable to expect that he will not walk very far. However, if he can save considerable amounts (as in the second example), presumably he will walk farther. In the final models of the analysis, the slope of the parking cost profile was taken as the parking cost difference between the distance of the first and third blocks.

Duration of Parking

It may be expected that, when duration of parking is shorter, people will park closer to their destinations (e.g., when making short business calls, keeping appointments, and so forth).

Income

With higher incomes, people may become more indifferent to parking costs. Dummy variables were used for income.

Age

Older people may be expected to walk shorter distances; walking may become very inconvenient for the elderly.

Sex

It may be that females walk less than males and respond differently when choosing parking locations. This variable was also a dummy.

Other Factors

In the analysis, parking lot capacity was not considered because it should have little effect on the choice process. The reason for this is that market prices will arrange themselves so that parking lots will be almost full without exceeding their capacities. If the capacity is exceeded or if the demand increases, then by the basic rules of economics prices will rise. This creates a balance, and capacities are not exceeded. Public parking facilities are sometimes underpriced and, therefore, they become congested. In general, however, their role in the overall structure of prices is not very important. Furthermore, the congestion itself has the effect of making the price higher. In the data set analyzed in this study, underpricing of public parking facilities was not an important factor.

Model Input

An example of the actual input to the analytical models developed is shown in Figure 2. It should be noted that the choice variable shown in Figure 2 (which is the dependent variable in the logit models) is one for the distance at which parking is accepted and for all greater distances. For shorter distances, it is 0. The models therefore give cumulative probabilities. Clearly, in the longest distance category, all

Individual Number	Actual Walking Time(sec.)	Actual Parking Cost(in cents)	Cheapest Parking Cost at Block						No. of Park. Hrs.	Age	Sex 0 = Male 1 = Female	Choice at the Block 1 = taken 0 = not taken						Income Group Dummy Variables			
			≤1	>1 ≤2	>2 ≤3	>3 ≤4	>4 ≤6	>6 ≤8				1	2	3	4	6	8	3	4	5	6
XXXX	350	175	325	275	225	175	125	50	8.0	40	0	0	0	1	1	1	1	0	0	1	0

Figure 2. Form of data input.

individuals have a choice variable of one. This indicates that all have accepted parking at or before that distance. The logit function is of the following form:

$$G(X)_{i^{th}\,block} = a_0 + a_1 \text{ (parking cost at } i^{th} \text{ block)} + a_2 \text{ (slope of the parking cost profile)} + a_3 \text{ (age)} + a_4 \text{ (sex)} + \sum_{i=5}^{8} a_i \text{ (income dummies)} + a_9 \text{ (number of parking hours)}$$

PARKING BEHAVIOR OVERVIEW

An initial picture of downtown parking behavior can most readily be gained by viewing a map showing parking costs, trip destinations, and actual parking places chosen. Figure 3 shows parking isocost lines in downtown Chicago derived from actual parking rates charged at the time of the northwest corridor travel survey. The map also shows the trip destinations of all the sampled persons who drove to downtown Chicago. For persons who walked a block or more from parking to their destinations, arrows extend from the trip destinations to the parking locations.

For those people who did walk, there was a definite pattern of parking and walking from lower parking cost areas to higher cost destinations. There is no obvious difference in distances walked from 1 part of downtown to another. The proportion of persons choosing to walk is higher where parking is expensive.

ANALYSIS OF WORK TRIPS

The general patterns of parking behavior shown in Figure 3 can be broken down in various ways and shown more precisely in tabular and graphical form. For work trips, Table 1 gives the number of drivers by distance of chosen parking from the trip destination. The majority of drivers (61.1 percent) park within 1 block of their destinations. The remaining 38.9 percent are spread throughout the remaining distances, the percentage of people parking at successively greater distances decreasing approximately linearly. The maximum distance for work trips is 6 blocks; in the sampled population nobody walked farther.

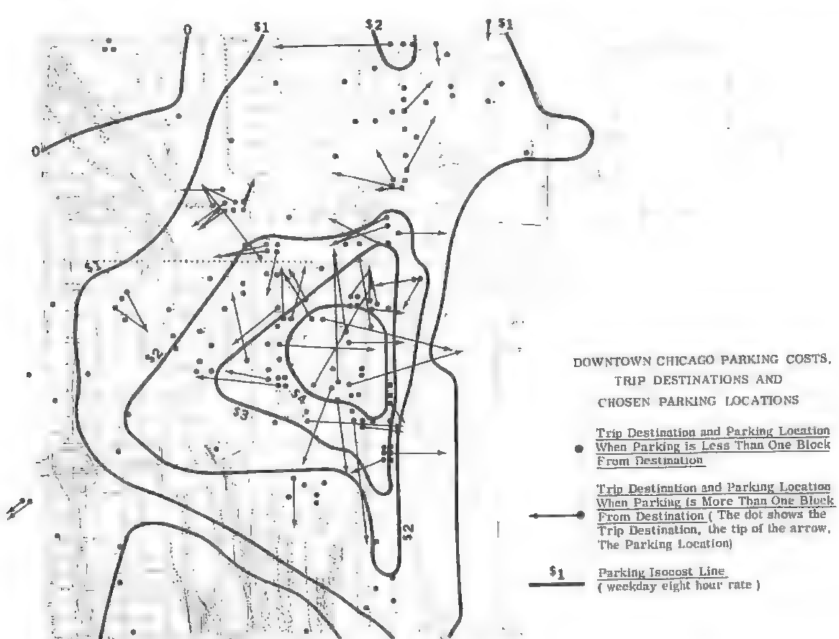


Figure 3. Locations in downtown Chicago of parking costs, trip destinations, and chosen parking locations.

TABLE 1

DISTRIBUTION OF PEOPLE AMONG PARKING PLACES FOR WORK TRIPS

Distance (1/2-mile blocks)	Drivers		
	Number	Percent	Cumulative Percent
> 0 and ≤ 1	69	61.1	61.1
> 1 and ≤ 2	17	15.0	76.1
> 2 and ≤ 3	14	12.4	88.5
> 3 and ≤ 4	8	7.0	95.5
> 4 and ≤ 6	5	4.5	100.0
> 6 and ≤ 8	0	0	
Total	113	100.0	

Parking costs, average parking costs, and average walking time are given in Table 2. There was only 1 person from the $5,000-to-7,999 income group who drove downtown, and thus data from this group are not included in Table 2. The threshold income group is therefore the $8,000-to-11,999 group. Average parking cost for this group is very low (68 cents), implying that people in this income range will not park downtown unless they can do so at low cost. The average cost of the next group ($12,000 to 16,999) is almost twice ($1.30) that of the preceding group. Subsequently, there is a small increase in the average costs paid by higher income groups. As overall averages, drivers walk 1.34 (1/12-mile) blocks and pay $1.34.

Relationships between walking time and other variables are shown by scatter diagrams. There was no observable relationship between the dependent variable (walking time) and variables of income and age. The scatter diagrams for first parking cost and parking cost difference are shown in Figures 4 and 5. Although variances are quite wide, there is, as expected, a relationship between these variables and walking time. Figure 4 shows that walking time increases with increasing parking cost in the first block, and Figure 5 shows that walking time increases with greater parking cost differences.

There were only 5 females among the 113 drivers in the sample. Because this variable could not be significant in the model, it was dropped.

ANALYSIS OF NONWORK TRIPS

As seen from data given in Table 3, a considerable portion of drivers making nonwork trips (42.3 percent) also park within 1 block. Unlike drivers on work trips, however, some drivers in this group park more than 6 blocks from their destinations (7.7 percent.

Of 52 nonwork drivers, 31 are male and 21 are female. On the average, males walk 229 sec and females walk 205 sec, with the overall average being 219 sec. This is substantially different from average walking time for work trips (134 sec) and corresponds approximately to 2.2 blocks.

The parking costs for different income groups are given in Table 4. Contrary to work trips, there are drivers in the lower income groups who make nonwork trips. The overall average cost paid ($1.38) is approximately equal to that of work trips ($1.34).

Similar scatter diagrams (walking time versus various other variables) were drawn for nonwork trips as for work trips. There were, however, no observable relationships

TABLE 2

PARKING COSTS BY INCOME GROUP FOR WORK TRIPS

Parking Cost[a]	Drivers										
	Income Group										
	$8,000 to 11,999		$12,000 to 16,999		$17,000 to 24,999		$25,000		Number	Percent	Cumulative Percent
	No.	Percent	No.	Percent	No.	Percent	No.	Percent			
$0.00 to 0.49	7	41.1	5	12.2	2	5.1	0	0.0	14	12.5	12.5
$0.50 to 0.99	2	11.8	4	12.2	2	5.1	3	20.0	12	10.7	23.2
$1.00 to 1.49	4	23.5	11	14.9	9	23.1	0	0.0	24	21.4	44.6
$1.50 to 1.99	4	23.5	14	33.2	15	38.5	6	40.0	39	34.8	79.4
$2.00 to 2.49	-	-	2	4.9	7	17.9	3	20.0	12	10.7	90.1
$2.50 to 2.99	-	-	2	4.9	3	7.7	3	20.0	8	7.2	97.3
Over $3.00	-	-	2	4.9	1	2.6	0	0.0	3	2.7	100.0
Total	17	15.2	41	36.6	39	34.8	15	13.4	112	100.0	
Avg parking cost, dollars	0.68		1.30		1.55		1.70		1.34		
Avg walking time, sec	128		120		155		128		134		

[a]Excluding company-paid trips.

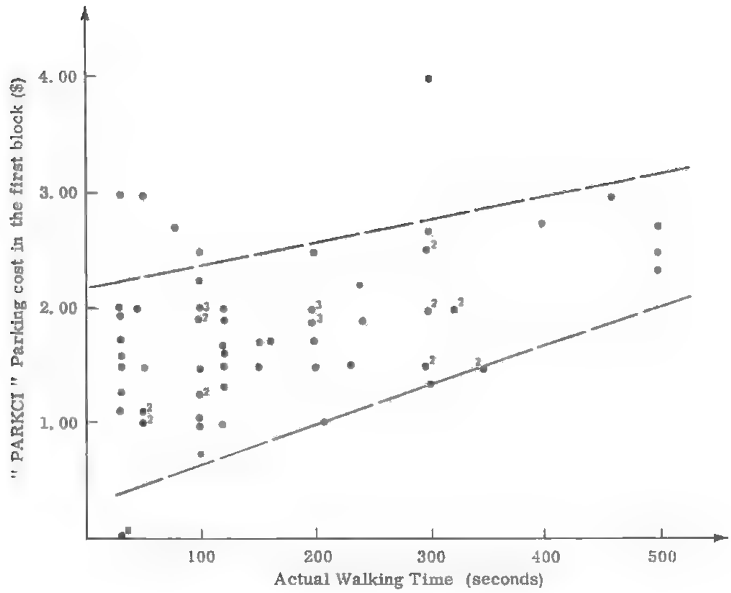


Figure 4. Walking time versus parking cost in first block.

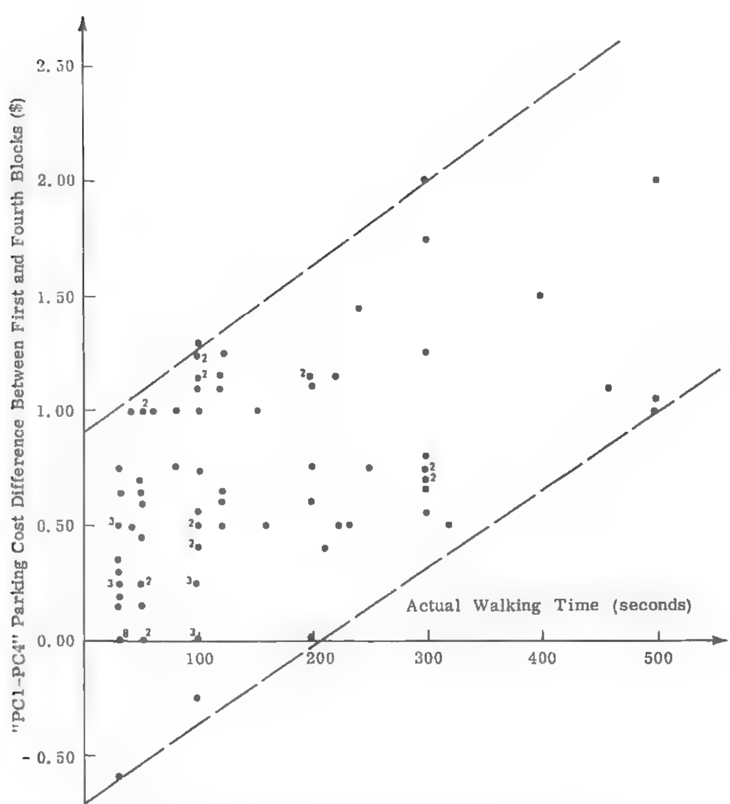


Figure 5. Walking time versus parking cost difference.

TABLE 3

DISTRIBUTION OF PEOPLE AMONG PARKING PLACES FOR NONWORK TRIPS

Distance (1/2-mile blocks)	Drivers		
	Number	Percent	Cumulative Percent
>0 and ≤ 1	22	42.3	42.3
>1 and ≤ 2	12	23.0	65.3
>2 and ≤ 3	8	15.4	80.7
>3 and ≤ 4	3	5.8	86.5
>4 and ≤ 6	3	5.8	92.3
>6 and ≤ 8	4	7.7	100.0
Total	52	100.0	

between the variables. Seemingly, people making nonwork trips walk longer distances while paying similar amounts, but their choices do not fall into patterns as clear as those observed in work trips. This may be because the knowledge of nonworkers about parking availability and costs is limited.

Because of these factors, the small sample size (52), and the very random nature of the scatter diagrams, none of the variables might have been expected to be significant in logit regression models. In spite of this, some logit models were tried. The results were insignificant.

CALIBRATION OF LOGIT MODELS FOR WORK TRIPS

As stated previously, a model for each distance was built. However, because in the work trip sample all persons had accepted a parking place by the sixth block, all of the choice variables were one for the distance between the fourth and sixth blocks. Therefore, only 4 models were built for the first 4 distances: one each for ≤ 1 block, for ≤ 2 blocks, for ≤ 3 blocks, and for ≤ 4 blocks. Because these models give the cumulative probabilities (i.e., the choice variable was coded as 1 for the distance at which parking was accepted and for all greater distances), the individual probabilities are as follows:

$$P(4 < X < 6) = 1 - P(X \leq 4)$$
$$P(3 < X \leq 4) = P(X \leq 4) - P(X \leq 3)$$
$$P(2 < X \leq 3) = P(X \leq 3) - P(X \leq 2)$$
$$P(1 < X \leq 2) = P(X \leq 2) - P(X \leq 1)$$
$$P(X \leq 1) = P(X \leq 1)$$

where $P(i < X \leq j)$ is the probability of parking at less than or equal to j blocks and more than i blocks. Hence, although there were 4 models, the $P(4 < X \leq 6)$ could also be found as given previously.

The models and associated statistics are given in Table 5. Successful work-trip models were developed for all distances up to 4 blocks from the destination. In all of the models, parking cost, parking cost difference, and distance affected the behavior of individuals in determining their likelihood of choosing a parking location. Also, in all models, the independent variable coefficients had the right sign and were of reasonable magnitudes. In all cases, the higher the parking cost was, the smaller the

TABLE 4

PARKING COSTS BY INCOME GROUP FOR NONWORK TRIPS

Parking Cost[a]	Drivers														
	Income Group												Number	Percent	Cumulative Percent
	$5,000		$5,000 to 7,999		$8,000 to 11,999		$12,000 to 16,999		$17,000 to 24,999		$25,000				
	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent			
$0.00 to 0.49	-	-	-	-	-	-	-	-	1	6.3	1	12.5	2	3.8	3.8
$0.50 to 0.99	-	-	1	33.3	2	18.2	4	33.3	4	25.0	1	12.5	12	23.1	26.9
$1.00 to 1.49	1	50	2	67.7	2	18.2	4	33.3	5	31.2	3	37.5	17	32.7	59.6
$1.50 to 1.99	1	50	-	-	6	54.5	-	-	2	12.5	1	12.5	10	19.2	78.8
$2.00 to 2.49	-	-	-	-	1	9.1	4	33.3	2	12.5	1	12.5	8	15.4	94.2
$2.50 to 2.99	-	-	-	-	-	-	-	-	-	-	-	-	0	0.0	94.2
$3.00 +	-	-	-	-	-	-	-	-	2	12.5	1	12.5	3	5.8	100.0
Total	2	3.8	3	5.8	11	21.1	12	23.1	16	30.8	8	15.4	52	100.0	
Avg parking cost, dollars	1.37		1.03		1.35		1.28		1.51		1.43		1.38		
Avg walking time, sec	140		420		172		227		186		280		219		

[a]Excluding company-paid trips.

TABLE 5

SUMMARY OF LOGIT MODELS FOR WORK TRIPS

Model	Variable	Estimate	Standard Error	t-Ratio	Likelihood Ratio[a]	Pseudo R^2
1 block	PARKC1	-0.00745	0.00385	1.93	25.76	0.276
	PC1-PC3	-0.01526	0.00672	2.26		
	CONSTANT	2.2571	0.55919	4.03		
2 blocks	PARKC2	-0.01099	0.004156	2.64	23.82	0.285
	PC1-PC3	-0.01444	0.005654	2.55		
	CONSTANT	3.4553	0.69	5.00		
3 blocks	PARKC3	-0.01504	0.00549	2.74	10.13	0.168
	PC1-PC3	-0.00943	0.00552	1.71		
	CONSTANT	4.3893	0.94553	4.64		
4 blocks	PARKC4	-0.01519	0.00872	1.71	6.451	0.182
	PCI-PC3	-0.01454	0.00733	1.98		
	CONSTANT	5.5711	1.3202	4.22		
6 blocks[b]						

[a]With 2 degrees of freedom.

[b]No model was built since P (X 6) = 1.

probability was of an individual choosing to park at the given distance. Similarly, the greater the parking cost difference was between the first and third block, the more likely people were to walk farther. Third, the simple effect of distance, as seen in the constant term, was that the greater the distance was, the less likely people were to walk that distance.

In addition, the sensitivity of the models to the parking cost and parking cost difference variables decreases as the distance increases (Table 5). A large portion of the probability is derived from the constant term. When other variables are 0, the constants give a probability of 0.9053 for the first model, 0.9694 for the second, 0.9878 for the third, and 0.9916 for the fourth. This makes sense because, when parking costs are 0, people can be expected not to walk. It is also seen that the first 2 models are more sensitive to parking cost differences than to parking costs. The sensitivity to these variables becomes approximately equal in the third and fourth models.

A number of tests were conducted to assess the statistical validity of the models. The results of several of the tests are given in Table 5. A brief discussion of each of the tests and their results are included in the following sections.

t-Test

The t-test is a test for significance of the coefficients. As seen from data given in Table 5, all coefficients are significant at the 0.10 level because all of the t-ratios are greater than the table values.

Likelihood-Ratio Test

To test for significance of the logit G(X) function, the likelihood ratio test was performed. This test works by proposing a null hypothesis that the probability, p_i, of an individual accepting parking at a given location is independent of the value of the parameters in the G(X) function. If this is true, the coefficients are 0. It can be shown that 2 times the log of the probability is approximately distributed as a χ^2 distribution with as many degrees of freedom as the number of variables in the equation (3). The value of χ^2 for 2 degrees of freedom (number of variables in the equation) at 0.05 level is 5.991. Because all the likelihood ratios are greater than this value, the relationships are significant at 0.05 level.

Pseudo-R^2 Test

The pseudo-R^2 test was developed mainly for discriminant analysis with a regression analogy. This measure tends to understate the effectiveness of the technique. The various statistics given in Table 5 give an overall and relative notion about the comparative efficacy of the 4 models.

APPLICATION OF THE FINAL MODEL: EFFECTS OF CHANGES IN PARKING PRICE POLICY

The effects of a particular type of change in downtown parking pricing policy were observed by changing all of the parking prices by given factors and by calculating the resultant expected proportions of persons parking at different locations, average expected costs, and average expected walking times. Table 6 gives the relevant statistics.

The percentages of people accepting the parking place and parking cost change factors were used to draw Figure 6. From the curves shown in Figure 6, the distribution of people at different locations, with a given parking price policy, can easily be found by drawing a vertical line at the desired factor and reading the percentages from the intersection of this vertical and the curves.

It is interesting to note that, if the original parking costs were to be increased proportionately about 1.5 times, 45 percent of the people would park in the first block and the remainder would be almost evenly distributed among the other block distances, as shown in Figure 6.

VALUE OF WALKING TIME FROM PARKING PLACE TO FINAL DOWNTOWN DESTINATION

Value of time may be expressed as an amount of money per unit of time (dollars/hour or cents/minute). As long as people can save at a rate greater than their value of time, they will walk from parking to destination; if they cannot, they will prefer not to walk. Thus, when the rate of savings becomes more or less equal to the value of time, people become indifferent to walking. Some will walk whereas others will not, the split presumably being 50:50. Based on the concept of indifference, a method was used for finding the value of walking time.

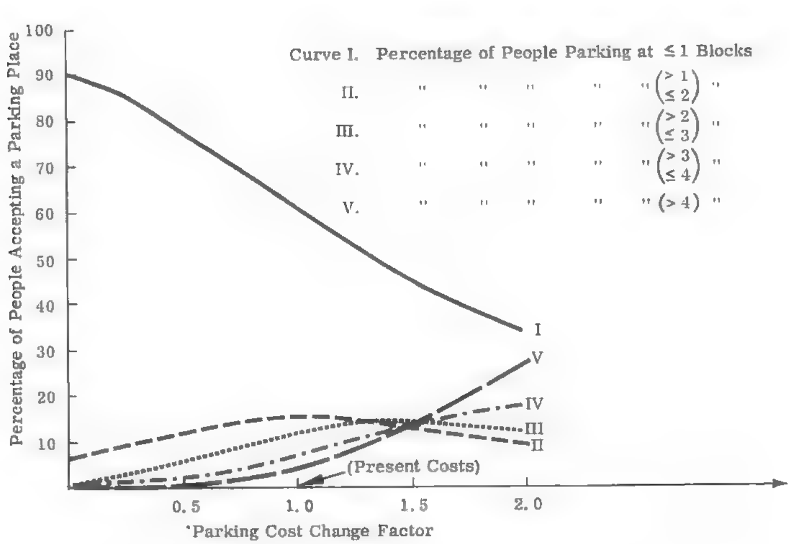


Figure 6. Effects of proportional changes in parking costs on distribution of people parking at different distances from destinations.

TABLE 6

EFFECTS OF CHANGES IN PARKING PRICE POLICY: WORK TRIPS

Parking Cost Change Factor	Average Expected Cost	Average Expected Walking Time	Expected Number of People at Blocks									
			1		1 and 2		2 and 3		3 and 4		4	
			No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent
0.0	0.0	64.3	102.3	90.5	7.3	6.4	2.1	1.8	1.0	0.8	0.4	0.4
0.3	41.80	74.61	95.5	84.5	11.0	9.7	3.9	3.5	1.8	1.6	0.8	0.7
0.5	67.39	85.81	88.9	78.6	13.8	12.2	6.22	5.5	2.77	2.4	1.3	1.2
0.8	100.93	110.36	76.9	68.1	16.6	14.7	11.1	9.8	5.4	4.8	2.96	2.6
1.0[a]	119.2	131	69	61.1	17	15.0	14	12.4	8	7.1	5	4.4
1.1	126.7	141.6	65.7	58.1	16.8	14.8	14.7	13.0	9.5	8.4	6.4	5.7
1.3	141.4	166.9	58.0	51.3	15.9	14.1	16.3	14.4	12.7	11.2	10.1	9.0
1.5	153.15	190.94	52.3	46.3	14.5	12.8	15.9	14.1	15.5	13.7	14.8	13.1
1.7	163.76	216.54	46.8	41.4	12.9	11.4	15.2	13.5	17.7	15.7	20.4	18.1
2.0	177.33	255.91	39.1	34.6	10.1	8.9	14.1	12.5	19.3	17.1	30.3	26.8

[a]Present.

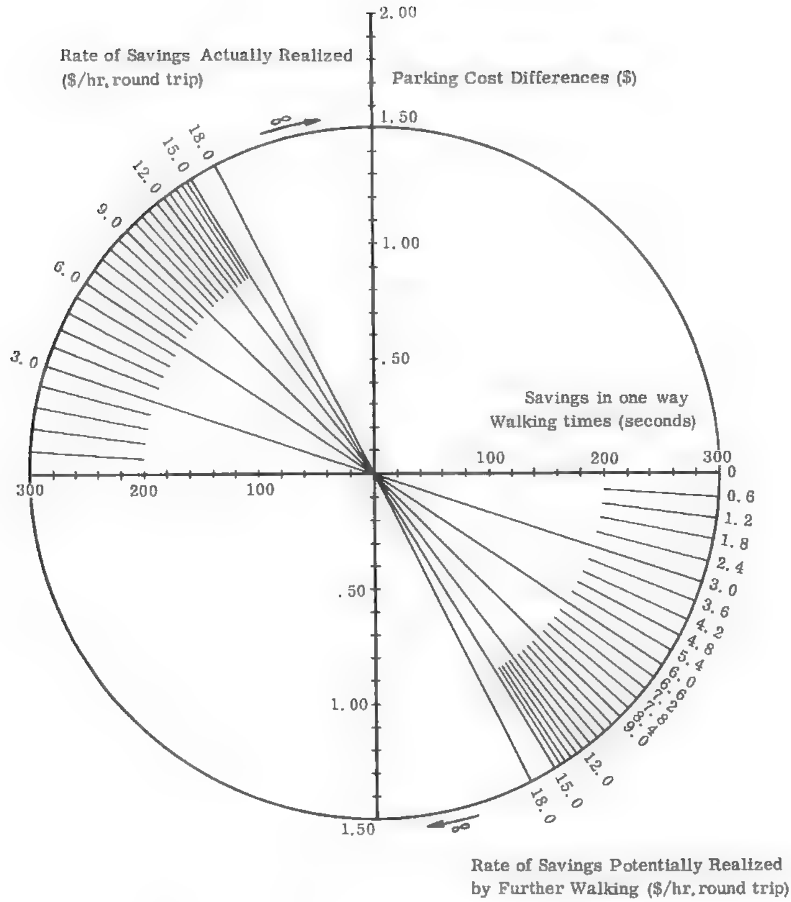


Figure 7. Walking time value indifference circle.

The tool used for finding the value of walking time is a circle, called a walking time value indifference circle, as shown in Figure 7 (11). When a person prefers not to walk to save time (even though he could save money by walking), a line, the slope of which represents the maximum value that saving time could have to him, is drawn in the fourth quadrant. The values around this portion of the circle represent the rates of savings that are potentially realized by further walking. However, if he actually walks to save money, a line is drawn in the second quadrant. In this portion of the circle, the values represent the rate of savings actually realized and are a minimum value of walking time for that individual.

In theory, if the value of time is a constant for all persons, the lower bound of "rate of savings actually realized" and the upper bound of "rate of savings potentially realized" should lie on the same line. The slope of this line will represent the value of walking time (Fig. 8) and will be called the "indifference line". In this case there should be no lines below the indifference line in either the second or the fourth quadrant. To the extent that different persons have different time values, there will be individuals with actual or potential savings below the indifference line.

This method was applied to the work trip data in the following manner. It was earlier established in this study that people are biased in favor of parking within 1 block of their destinations. For people who parked in this block, only the lines in the fourth quadrant representing rates of potentially realizable savings could be drawn because these parkers do not walk to save on parking cost. For these persons,

the maximum rate of savings available within a 6-block radius was used. The potential savings available to these persons (Fig. 9) ranged from 0 to a maximum of $11.50/hour. The mean rejected saving was about $5/hour. The fact that a considerable fraction of the total sample rejected possible savings ranging from $5 to $11.50/hour walking is an indication of the bias for parking within the first block.

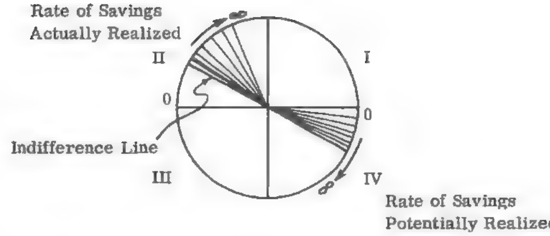


Figure 8. Walking time value indifference circle: theoretical data plot.

A second analysis was performed on the drivers who parked at more than the 1-block distance. For these people, the value of time was found by dividing the money saved by the time spent, and a line was drawn in the second quadrant. Also, because many of these people were subject to still greater savings by walking farther, a line was drawn in the fourth quadrant to represent the further potential savings that they did not choose to gain. Because of the difficulties of establishing reliable indifference lines for individual block distances, an indifference circle, including all individuals who parked more than 1 block from their destinations was drawn (Fig. 10). A clear indifference line for all except only a few people was established at a value of time of $4.50/hour, which represents the value of time for roughly 40 percent of the population (those persons who did choose to walk). As indicated previously, for the approximately 60 percent of the population who did not choose to walk more than a block, higher time values could be derived due to the bias in favor of the first block.

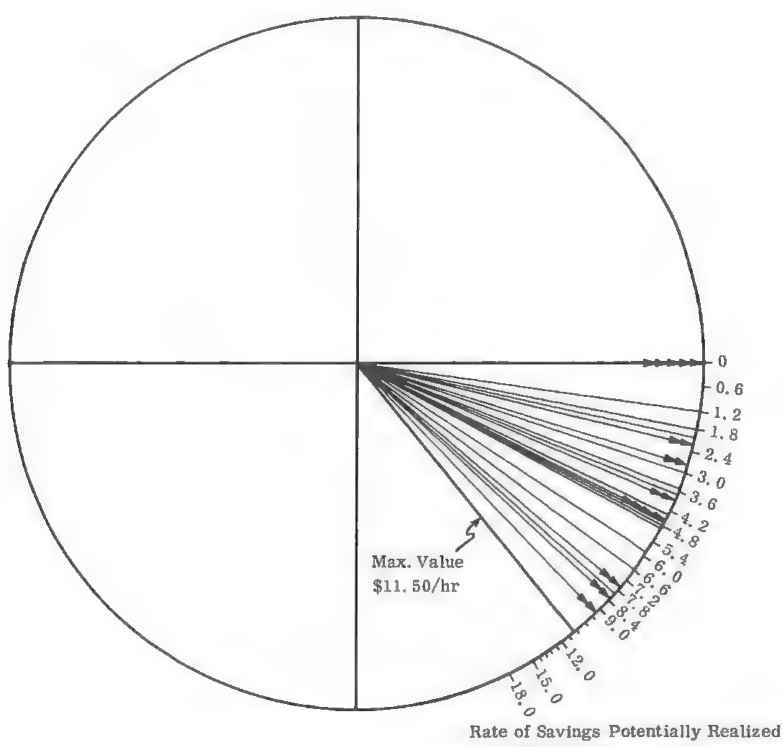


Figure 9. Walking time value indifference circle for first block.

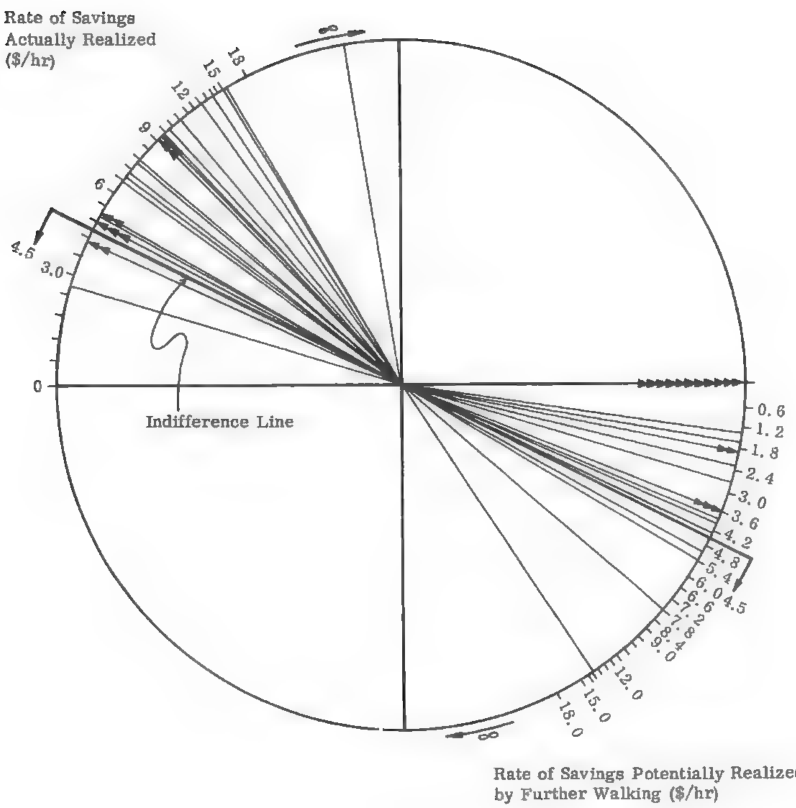


Figure 10. Walking time value indifference circle for more than one block.

CONCLUSIONS

It was observed that drivers have a strong bias for parking in the first block for work trips to Chicago's downtown area. It was also found that commuters are sensitive in their parking and walking behavior to absolute parking costs and to available savings through walking. The higher the parking costs are at the destination, the far-

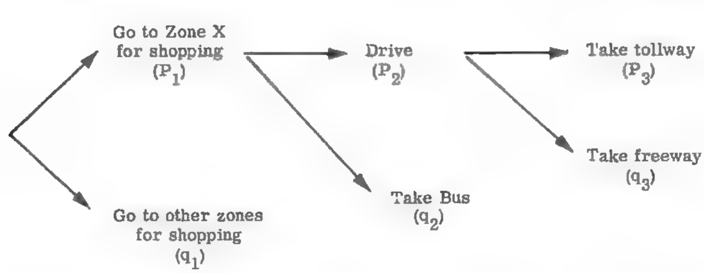


Figure 11. Example of multiplicative probabilities for multinomial mode choices.

ther people are likely to walk. Also, the greater the available savings are through walking, the more likely commuters are to walk farther.

In the models, income was found to be insignificant. This is apparently because its effect is largely taken care of in the overall travel mode choice. To make decisions to drive downtown, people first must consider parking costs and walking distances and make comparisons with their incomes. Once the decision to drive downtown is made, income apparently becomes an insignificant variable.

Another reason for the insignificance of the socioeconomic factors may be that northwest corridor modal-split data were biased toward high-income groups. Even in the transit user sample, there were very few people having incomes less than $8,000.

The value of time was found to be $4.50/hour for persons who did walk from parking places. However, this value of time is only for persons with incomes over $8,000. This should be kept in mind when this figure is applied for cost-benefit analysis.

The final probabilistic model was virtually multinomial. Therefore, this approach may well be applied to other problems requiring a multinomial solution. In parking problems, the probabilities are additive. However, this approach can be applied to problems where the probabilities are multiplicative as well. An example of such a situation is given in Figure 11.

The probability that a person will drive to zone X via a tollway will be $p = p_1p_2p_3$. Notice that the only requirement is that the choices are binary at each stage. For this example, models may be built for the 3 stages given and then combined to give cumulative probability. This approach has numerous applications toward finding solutions to transportation modal-choice problems.

REFERENCES

1. Beesley, M. E. The Value of Time Spent 24th May 71. Economica, May 1965.
2. Finney, D. J. Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve, 2nd Ed. Cambridge Univ. Press, England, 1962.
3. Draper, N. R., Smith, H. Applied Regression Analysis. John Wiley and Sons, New York, 1966.
4. Ehrenfeld, S., and Sebastion, L. Introduction to Statistical Method. McGraw-Hill, New York, 1964.
5. Lisco, T. E. The Value of Commuters' Travel Time: A Study in Urban Transportation. Univ. of Chicago, PhD dissertation, 1967.
6. Lisco, T. E. Project Statement: Northwest Chicago Corridor Modal Split Project. Chicago Area Transportation Study, 1969.
7. Lisco, T. E., and Stopher, P. R. Modelling Travel Demand: A Disaggregate Behavioral Approach—Issues and Applications. Proc., Eleventh Annual Meeting of Transportation Research Forum, 1970, pp. 195-214.
8. Stopher, P. R. A Probability Model of Travel Mode Choice for the Work Journey. Highway Research Record 283, 1969, pp. 57-65.
9. Stopher, P. R. Course Notes of Transportation Analysis Methods, Northwestern Univ., Evanston, Winter 1970.
10. Theil, H. A Multinomial Extension of the Linear Logit Model. Center for Mathematical Studies in Business and Economics, Univ. of Chicago, Report 6631, 1966.
11. Warner, S. L. Stochastic Choice of Mode in Urban Travel. Northwestern Univ. Press, 1962.

THE n-DIMENSIONAL LOGIT MODEL: DEVELOPMENT AND APPLICATION

Paul R. Rassam, Raymond H. Ellis, and John C. Bennett,
Peat, Marwick, Mitchell and Company, Washington, D.C.

This paper discusses the development, calibration, and testing of the n-dimensional logit model, a modal-split model that appears to appropriately respond to the n-mode situation and that, at least conceptually, addresses the new mode problem. The model is based on the assumption that the ratio of a small change in modal split of a given mode to that of a given transportation variable is proportional to the modal split of this mode and to a linear function of the modal split of all modes. In conjunction with the mathematical definition relating to modal split, these assumptions lead to a system of differential equations defining modal-market shares as functions of the transportation attributes. This formulation does not require that the same set of variables be used to define the transportation attributes of each of the modes. The model was applied to predict the market shares of 4 modes—private automobile, rental car, taxi, and limousine—used for access (or egress) to (or from) airports in the Baltimore-Washington area. Calibration results indicate that the coefficients estimated have the expected signs and relative magnitudes as well as low standard errors of estimate. Testing indicates that the model displays the appropriate sensitivities and adequately reproduces the aggregate trips for each mode. Overall, results of this initial work are sufficiently promising that further application of the model in both operational and experimental situations appears warranted.

•THE INCREASE IN RECOGNITION of the dollar and social costs associated with programs mainly advocating the construction of urban freeways have led planners and decision-makers to seek other strategies for alleviating the urban transportation problem. Many urban transportation concepts, ranging widely in the extent of technological and institutional innovation required for implementation, have been proposed in recent years.

However, it is not at all clear that current modal-split models are capable of adequately predicting passenger demands for these new transportation systems. Aside from the difficulties inherent to introducing a new mode in a model structure, many of the existing modal-split models are designed to deal with dual choices between the 2 current urban transportation modes, automobile and public transit. This limitation obliges the analyst to lump into the public-transit category 2 or 3 modes with widely differing characteristics. It requires the a priori definition of aggregate transit impedances and, in most cases, the breakdown of the nonautomobile market into specific transit categories.

In view of the preceding, it appears desirable to develop a modal-split model that has the capability to predict simultaneously the market shares of 2 or more modes operating in a competitive environment. This paper discusses the development, calibration, and testing of the n-dimensional logit model, a modal-split model that appears to appropriately respond to the n-mode problem and that, at least conceptually, addresses the new mode problem.

Sponsored by Task Force on Value of Travel Time.

The model discussed in this paper belongs to a general class of models known as stimulus-response models (1). Such models have been previously used in transportation studies. Two early examples of stochastic disaggregate modal-choice models are those of Lisco (2) and Stopher (3). Lisco used probit analysis to derive the value of time for commuters in the Chicago area. Stopher used logit analysis to model the modal choice of work trips in London, England. In a recent paper, Stopher and Reichman (4) have given an excellent discussion of the modal-split problem, focusing particularly on the potential of disaggregate stochastic modal-choice models.

MODEL DEVELOPMENT

Let w and x identify respectively the dependent variables (modal splits) and the independent variables (transportation or eventually dummy variables). Let m be an index identifying a given mode ($m = 1, \ldots, j, \ldots, k, \ldots, M$) and i be an index identifying a given transportation attribute, e.g., time or cost ($i = 1, \ldots, I$). Hence, the share of mode m is noted w_m and the ith attribute of mode j is noted x_{ij}.

Two types of assumptions are required. The first assumption pertains to the general nature of any modal-split problem: (a) The modal split of each mode included is between 0 and 1, and their sum is equal to unity; (b) modal splits are monotonic functions of the independent variables; and (c) if the transportation variables are expressed in units such that the disutility of traveling by a given mode is an increasing (decreasing) function of its transportation variables, then the share of that mode decreases (increases) when any of its transportation variables increases (decreases) and, all other things equal, those of the other modes increase (decrease). The second pertains to the specific premises postulated to structure the relationship between modal split and the explanatory transportation variables; that is, the ratio of a small change in modal split of a given mode to that of a given transportation variable is proportional to the modal split of this mode and to a linear (homogeneous) function of the modal splits of all modes. If it is assumed that these are the requisite continuity and derivability conditions, this relationship is expressed mathematically as

$$\frac{\partial w_m}{\partial x_{ij}} = w_m \sum_k \alpha_{ijmk} w_k \tag{1}$$

where α_{ijmk} is a coefficient to be determined. A total of ($M \times I \times M$) such equations can be written for all possible permutations of m, i, and j. Together they constitute a system of differential equations defining the functions w_m. However, before these equations are integrated, the nature of the coefficients α_{ijmk} should be examined in the context of the assumptions formulated previously. It can be shown that several classes of coefficients α are, in fact, equal to 0 and that there exist very simple relationships among those that are different from 0. Specifically, the coefficients α are such that $\alpha_{ijkm} = 0$ if ($j \neq k \neq m$), $\alpha_{ijkj} = -\alpha_{ijjk}$ if ($j \neq k$), $\alpha_{ijkk} = 0$, and $\alpha_{ijjk} = \alpha_{ijjm} = \alpha_{ij}$. When the preceding constraints are considered, the underlying assumption of the model as stated by Eq. 1 becomes

$$\frac{\partial w_m}{\partial x_{ij}} = \Delta_{jm} w_m \tag{2}$$

where

$$\Delta_{jm} = \begin{cases} -\alpha_{ij} w_j & \text{if } (m \neq k) \\ \alpha_{im}(1 - w_m) & \text{if } (m = k) \end{cases}$$

It can be shown that the solution of the differential system defined by Eq. 2 is

$$w_m = \frac{\exp\left(\sum_i \alpha_{im} x_{im} + a_m\right)}{\sum_j \exp\left(\sum_i \alpha_{ij} x_{ij} + \alpha_j\right)} \tag{3}$$

where α_j is a mode-specific (integration) constant.

It should be noted that the sign of the coefficient α_{ij} depends on the nature of the variable x_{ij} to which it is attached. For example, if i corresponds to frequency, then α_{im} must be positive because the share of mode m increases as its frequency increases. Conversely, if i corresponds to travel time, α_{im} must be negative in order for the modal split of m to increase when its travel time is improved. These prerequisite sign conditions must be met by the coefficients derived from calibration. If this were not the case, the prediction capability of the model would be poor to begin with, inasmuch as decreases in the level of service of a given mode would result in higher market share for that mode. Finally, it should be observed that no assumption must be made as to the number of attributes considered for each mode. This feature introduces into the analysis a degree of flexibility lacking in models incorporating ratios or differences of variables.

Equation 2 can be used to derive the elasticity of a given w_m with respect to a given x_{ik}, as in the following:

$$\frac{x_{ik}}{w_m} \frac{\partial w_m}{\partial x_{ik}} = \Delta_{km} x_{ik} \tag{4}$$

In other words, Eq. 4 shows that the elasticity of w_m with respect to x_{ik}, a transportation variable of another mode k, is proportional to x_{ik} and to the share w_k of that mode; the elasticity of w_m with respect to one of its own transportation variables x_{im} is proportional to x_{im} and to the market share that mode m can still gain $(1 - w_m)$.

CALIBRATION

Two alternatives are available for calibrating the proposed modal-split model—simultaneous least squares and maximum likelihood. Because of the additive structure of the denominator of Eq. 3, the model does not lend itself to least squares calibration. However, if we single out 1 mode, say mode u, and call it a base mode, then the natural logarithm of the ratio of the modal split of any given mode m to that of mode u is a linear function of the independent variables. If there are M modes competing for a given market, then (M - 1) such equations have to be calibrated, one for each mode except mode u. This leads to (M - 1) sets of coefficients for mode u. Because these coefficients may vary substantially, it is suggested that, instead of minimizing the square error for each equation considered separately, the square error of all the equations considered simultaneously be minimized.

The second alternative involves maximizing the likelihood function of the logit function. In this case, contrary to the preceding, individual rather than aggregate observations are required. When only aggregate observations are available, they can be converted to individual observations if the number of trips is known. If not, and this would imply giving equal weights to each zone pair, a common trip number can be used.

More specifically, the quantity maximized is the logarithm of the likelihood function. The coefficients are determined by successive iterations, the first estimate of the likelihood being that of the multinomial model, that is to say, the average modal splits of the sample. The method allows establishing constraints between any subset of the coefficients. However, because of its iterative nature and, in some cases, the large number of observations to be processed, this method is more time-consuming than the simultaneous least squares method that requires summing variables and cross products thereof, followed by the solution of a system of linear equations.

MODEL APPLICATION

Trip Stratification

The n-dimensional logit model was calibrated with data collected by the Washington-Baltimore airport access survey. Essentially, 2 trip-purpose stratifications were considered, namely, business and nonbusiness trips. This initial stratification is dictated by the behavior of business travelers who emphasize access time and somewhat disregard cost, whereas the opposite is true for nonbusiness travelers. A further breakdown within this initial stratification is justified by the markedly different overall trip patterns generated by the origin end of a trip. As can be expected, the private automobile is an even more convenient mode for departing travelers than for arriving travelers. Table 1 gives average modal splits for the stratifications (i.e., business and nonbusiness travelers going from and to the airport and for 5 modes: private car, rented car, taxi, limousine, and public bus. The business trip market is shared, to a very large extent, by the first 4 modes inasmuch as bus trips account for less than 2 percent of the total trips in either direction. Because of this situation, only 4 modes were retained in the analysis; that is, private car, rented car, taxi, and limousine. Similarly, in the case of nonbusiness trips, the rented car, which accounts for about 4 percent of the total trips in both directions, was not included in the analysis and, hence, 3 modes were retained—private car, taxi, and limousine.

The modal-split analysis performed in this study is, to a large extent, comparable to those analyses performed in classic urban transportation studies. These studies have generally attached a great significance to the nature of the trip end, e.g., by distinguishing between trip ends to the home and to other destinations. In the present analysis, the corollary is whether a traveler resides in the area under study. This distinction is important inasmuch as it generally determines the availability of a private car. This important information was not collected by the study. Had it been available, the stratification by residents and nonresidents would have precluded, for most cases, the use of private car for the nonresident traveler. An attempt was made to supplement this deficiency by assuming that, in all likelihood, most nonresident travelers would originate in or be destined to the central business districts and, therefore, that trips should be further broken down by central business districts and residential districts. However, this approach did not improve the quality of the business-to-airport model to which it was applied.

Data Preparation

Impedance estimation is open to a considerable range of detail and specification, particularly with respect to perceived impedances vis-a-vis actual-time measures and cost estimates based on marginal or average cost models. Component measures of the travel service provided by each mode were derived from network analysis, fare and

TABLE 1

AIR TRAVELER MODAL SPLIT BY DIRECTION

Mode	Percentage of Airport Trips			
	Business		Nonbusiness	
	To	From	To	From
Private automobile	45.4	32.6	63.5	57.1
Rented automobile	9.2	10.5	3.7	4.3
Taxi	32.1	31.2	22.7	19.9
Limousine or coach	12.3	23.9	8.3	16.3
Bus	1.0	1.8	1.8	2.4

time schedules, ground counts at airports, and data reported and processed from the Baltimore-Washington airport access survey. The component measures were selectively combined to form impedance measures for each mode as shown in Figure 1. The set of impedance measure equations used in the calibration are shown in Figure 2. The effects of changes in component service measures (e.g., parking rates) on aggregate impedance measures can be explicitly determined by evaluating these equations.

A data set was constructed for the travel survey period in 1966. This consisted of travel volumes and average impedance measures for each mode between each of the 78 zones and each of the 3 airports. For each mode, the time components were divided into main modal and submodal times (e.g., limousine riding time and access to limousine time) and in-vehicle and out-of-vehicle times. Nine different time measures were available from this classification, as given in Table 2. Travel cost was simply broken down into submodal cost, main modal cost, and total cost. Impedances used for calibration are average daily values and are the same for both purposes.

An observation should be made concerning the nature of the time and cost impedances that, in many instances, were derived from distance measures. As might be expected, time and cost are correlated and the resulting collinearity is a source of problems because it increases, sometimes substantially, the standard error of estimate of the coefficients.

Model Calibration

Several prototype models, involving different combinations of time and cost variables, were tested initially. Specifically, business models were calibrated by maximum likelihood, with the following variables:

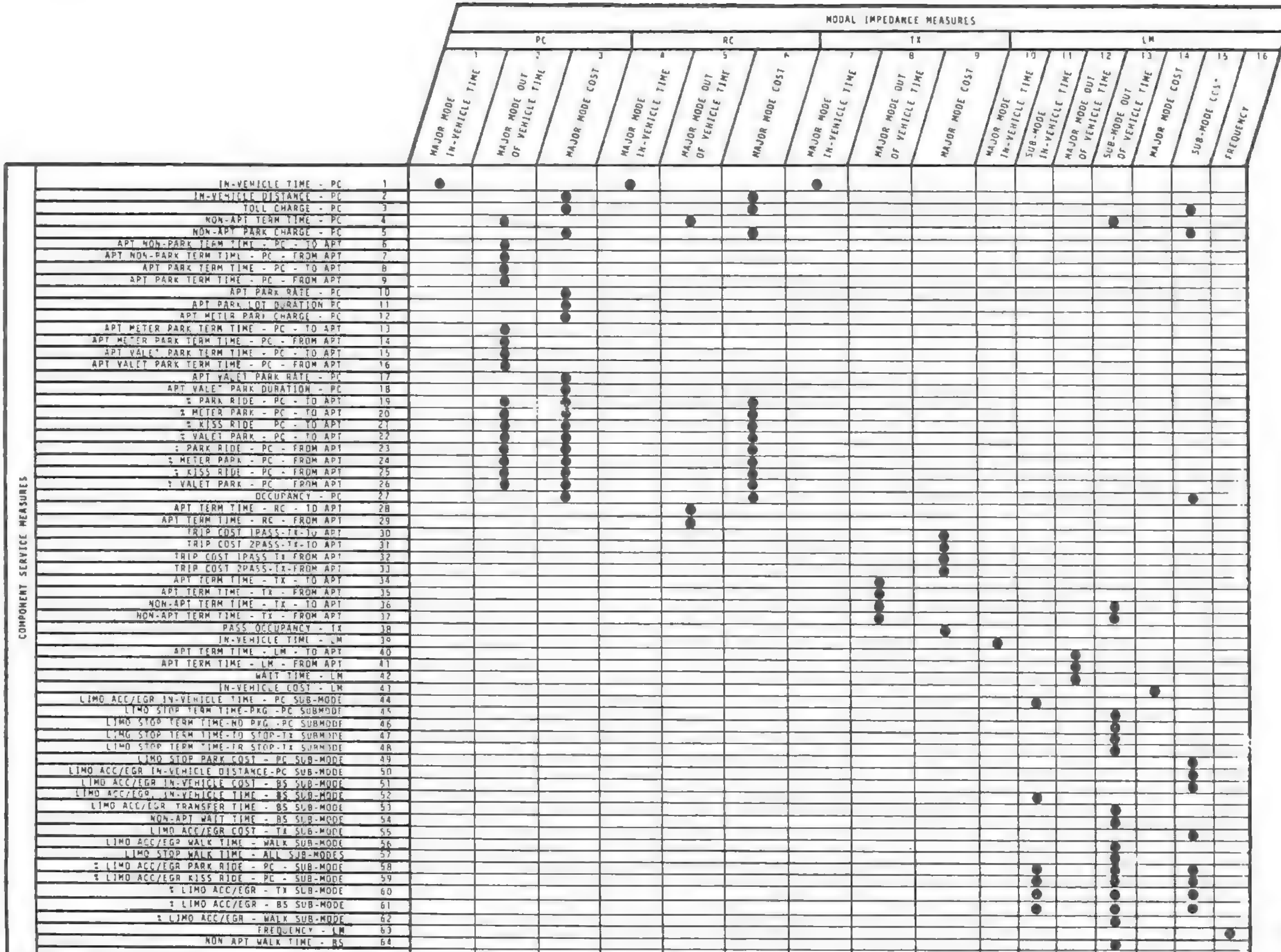
1. Total times and costs for all modes;
2. Total time and parking cost for the private automobile and total times and costs for the remaining 3 modes;
3. Total times and costs for all modes except limousine for which in-vehicle time, out-of-vehicle time, and total cost were incorporated; and
4. Total times and costs for all modes except limousine for which in-limousine time, out-of-limousine time, and total cost were incorporated.

These models were calibrated for each directional stratification. In most cases, the signs of the coefficients were acceptable. However, the standard errors were often large and even, in some instances, exceeded the magnitude of the corresponding coefficient. To overcome these problems, the directional stratification was abandoned and, as mentioned earlier, the sample was stratified by central business districts and residential districts. Neither approach provided acceptable results.

TABLE 2

TRIP TIME COMPONENTS

Travel Classification	Modal Classification		
	Submode	Main Mode	Total
In vehicle	Submode in-vehicle time	Main mode in-vehicle time	Total in-vehicle time
Out of vehicle	Submode out-of-vehicle time	Main mode out-of-vehicle time	Total out-of-vehicle time
Total	Total submode time	Total main mode time	Total travel time



PC = PRIVATE CAR
RC = RENTAL CAR
TX = TAXI
LM = LIMOUSINE
BS = BUS

Figure 1. Modal impedance measures.

$$Y_{1,t} = Y_{1,f} = X_1$$

$$Y_{2,t} = X_4 + (X_6)(X_{21}) + (X_8)(X_{19}) + (X_{13})(X_{20}) + (X_{15})(X_{22})$$

$$Y_{2,f} = X_4 + (X_7)(X_{25}) + (X_9)(X_{23}) + (X_{14})(X_{24}) + (X_{16})(X_{26})$$

$$Y_{3,t} = \left[(X_2)(\alpha_1) + X_3 + X_5 + (X_{10})(X_{11})(X_{19}) + (X_{17})(X_{18})(X_{22})\right]\left[\frac{X_{19} + X_{22}}{X_{27} - (X_{21} + X_{20})(2.0)}\right] + \left[(X_2)(\alpha_1) + X_3 + X_5\right](X_{20} + X_{21}) + (X_{12})(X_{20})$$

$$Y_{3,f} = \left[(X_2)(\alpha_1) + X_3 + X_5 + (X_{10})(X_{11})(X_{23}) + (X_{17})(X_{18})(X_{26})\right]\left[\frac{X_{23} + X_{26}}{X_{27} - (X_{25} + X_{24})(2.0)}\right] + \left[(X_2)(\alpha_1) + X_3 + X_5\right](X_{24} + X_{25}) + (X_{12})(X_{24})$$

$$Y_{4,t} = Y_{4,f} = X_1 \qquad Y_{5,t} = X_4 + X_{28} \qquad Y_{5,f} = X_4 + X_{29}$$

$$Y_{6,t} = \left[(X_2)(\alpha_2) + X_3 + X_5 + \alpha_3\right]\left[\frac{X_{19} + X_{22}}{X_{27} - (X_{21} + X_{20})(2.0)}\right]$$

$$Y_{6,f} = \left[(X_2)(\alpha_2) + X_3 + X_5 + \alpha_3\right]\left[\frac{X_{23} + X_{26}}{X_{27} - (X_{25} + X_{24})(2.0)}\right]$$

$$Y_{7,t} = Y_{7,f} = X_1 \qquad Y_{8,t} = X_{34} + X_{36} \qquad Y_{8,f} = X_{35} + X_{37}$$

$$Y_{9,t} = (X_{38} - 1.0)(X_{31})(0.5) + (2.0 - X_{38})(X_{30})$$

$$Y_{9,f} = (X_{38} - 1.0)(X_{33})(0.5) + (2.0 - X_{38})(X_{32}) \qquad Y_{10,t} = Y_{10,f} = X_{39}$$

$$Y_{11,t} = Y_{11,f} = (X_{44})(X_{58} + X_{59} + X_{60}) + (X_{52})(X_{61})$$

$$Y_{12,t} = X_{40} + X_{42} \qquad Y_{12,f} = X_{41} + X_{42}$$

$$Y_{13,t} = (X_4 + X_{45} + X_{57})(X_{58}) + (X_4 + X_{46} + X_{57})(X_{59}) + (X_{47} + X_{36} + X_{57})(X_{60}) + (X_{56} + X_{57})(X_{62}) + (X_{64} + X_{54} + X_{53} + X_{57})(X_{61})$$

$$Y_{13,f} = (X_4 + X_{45} + X_{57})(X_{58}) + (X_4 + X_{46} + X_{57})(X_{59}) + (X_{48} + X_{37} + X_{57})(X_{60}) + (X_{64} + X_{54} + X_{53} + X_{57})(X_{61}) + (X_{56} + X_{57})(X_{62})$$

$$Y_{14,t} = Y_{14,f} = X_{43}$$

$$Y_{15,t} = Y_{15,f} = \left[X_3 + X_5 + X_{49} + (X_{50})(\alpha_1)\right]\left[\frac{X_{58}}{X_{27} - (X_{59})(2)}\right] + \left[X_3 + X_5 + (X_{50})(\alpha_1)\right](X_{59}) + (X_{55})(X_{60}) + (X_{51})(X_{61})$$

$$Y_{16,t} = Y_{16,f} = X_{63}$$

x = component service measure

y = modal impedance measure

α_1 = cost per mile for private car

α_2 = cost per mile for rental car

α_3 = fixed charge rental car

t = to airport

f = from airport

subscripts refer to figure 2

Figure 2. Impedance measures formulas.

Total time and total cost were correlated and combined for each mode to yield a single impedance measure. Two series of calibration runs were performed for each of the 4 stratifications. In the first series, no occupancy was taken into account. In the second series, occupancies of 1.67 and 1.40 were assumed for automobile and taxi respectively. In all cases, the standard errors of estimate of the coefficients were low and in many instances were under 10 percent, and coefficients always displayed the expected signs. In addition, these coefficients were always significantly different from 0, far beyond the 0.001 level according to the maximum likelihood ratio test (minus twice the logarithm of the likelihood ratio), which is distributed as χ^2 with as many degrees of freedom as there are coefficients estimated.

The choice of the final model was determined by the coefficient magnitude that directly influences the sensitivity of the model. Both types of models display relatively high sensitivities, particularly in the case of private car cost. This condition is accentuated when occupancy is not taken into account. For this reason, it was decided to choose the model series incorporating occupancy. The results are given in Table 3 for which the following observations can be made.

For each mode and each trip purpose stratification, travelers seem to be more sensitive, and hence more likely to change modes, when going (in terms of access) to rather than when leaving the airport. Because the calibration sample does not distinguish between residents and nonresidents, this cannot, strictly speaking, be explained by familiarity with the region. Among other factors, it is due to the importance attached by the traveler to "making his flight".

The high sensitivity of the private car to cost is an indicator showing that cost, as expected, can be a powerful deterrent from using a car, especially for long-duration trips. Most probably, this would have appeared were this mode broken into 3 classes: those travelers who are driven to the airport, those who park for a short duration (2 days or less), and those who park for a long duration. As noted earlier, parking cost is an average measure based on the average proportion of travelers who park their vehicles at the airports and the average duration of parking.

Between the 2 purpose stratifications, it is surprising to note that the private car cost coefficients, and therefore the sensitivities to this measure, are higher in the case of business trips, regardless of the trip direction. This unexpected result could be attributed to the fact that nonbusiness travelers are often accompanied to or met at the airport. This reason also suggests that the higher cost coefficients of the business trips reflect the cost of the unavailability to the rest of the household of a personal vehicle parked at an airport.

Rented car sensitivity to time is low, and, in fact, ranks third following taxi and private car. This probably suggests that the businessman renting a car does not use it solely for airport access reasons. In most cases, a taxi ride would be more convenient and faster, in particular when downtown parking is required. It can be conjectured that an automobile has the advantage of "flexibility" when several trips are projected in the area visited.

Taxi displays, somewhat markedly and regardless of the trip purpose, the highest sensitivity to time. Two related justifications to this could be found. National Airport "points" dominate the calibration sample, and taxi fares to and from the airport are relatively low, which makes this mode particularly attractive.

The relatively low sensitivity to taxi cost can be explained by the latter justification presented in the preceding, i.e., the relatively low fare of taxi. Most probably, were this fare significantly higher and, hence, not within the means of most travelers, the sensitivity might be higher.

As expected, limousine displays the least sensitivity to time and second highest sensitivity to cost after private car. This is due to, in all likelihood, the competition offered by taxi that is faster and, in absolute terms, not much more expensive (especially in group riding) to areas close to an airport. Hence, were limousine cost to increase, all other things equal, in any significant amount, one would expect the attraction of taxi to be "irresistible".

The least squares techniques described previously were also used in the calibration runs. When applied to models in which total times and costs were introduced for

TABLE 3

CALIBRATED TIME AND COST COEFFICIENTS

Trip Purpose Stratification	Private Car		Rented Car		Taxi		Limousine	
	Time (min)	Cost (dollars)	Time (min)	Cost (dollars)	Time (min)	Cost (dollars)	Time (min)	Cost (dollars)
Business trips to airports	-0.0401	-1.9097 (0.125)	-0.0265	-0.6030 (0.030)	-0.0624	-0.3962 (0.024)	-0.0224	-1.0646 (0.057)
Business trips from airports	-0.0276	-1.3031 (0.075)	-0.0150	-0.3505 (0.015)	-0.0642	-0.3447 (0.020)	-0.0106	-0.5482 (0.030)
Nonbusiness trips to airports	-0.0342	-1.6058 (0.203)	—	—	-0.0631	-0.4025 (0.030)	-0.0177	-0.7922 (0.142)
Nonbusiness trips from airports	-0.0144	-0.6767 (0.163)	—	—	-0.0573	0.3082 (0.025)	-0.0104	-0.5348 (0.110)

TABLE 4

TEST RESULTS OF BUSINESS TRIPS TO AND FROM AIRPORTS

Mode	Direction	Unadjusted Results				Adjusted Results			
		Observed Mean Trips	Estimated Mean Trips	Correlation Coefficient	Root Mean Square Error	Observed Mean Trips	Estimated Mean Trips	Correlation Coefficient	Root Mean Square Error
Private car	To	103.1	102.4	0.96	100.6	103.1	103.4	0.96	61.3
	From	75.4	63.2	0.91	61.5	75.4	75.3	0.92	61.1
Rented car	To	19.2	22.4	0.88	20.2	19.2	19.1	0.88	18.4
	From	20.7	0.7	0.81	33.7	20.7	21.2	0.88	18.4
Taxi	To	83.3	73.8	0.99	86.2	83.2	83.0	0.99	68.8
	From	79.9	144.5	0.99	152.1	79.9	80.2	0.99	41.3
Limousine	To	31.5	38.4	0.87	53.6	31.5	31.4	0.88	48.2
	From	61.4	29.0	0.83	122.0	61.4	60.7	0.96	63.1

TABLE 5

TEST RESULTS OF NONBUSINESS TRIPS TO AND FROM AIRPORTS

		Unadjusted Results				Adjusted Results			
Mode	Direction	Observed Mean Trips	Estimated Mean Trips	Correlation Coefficient	Root Mean Square Error	Observed Mean Trips	Estimated Mean Trips	Correlation Coefficient	Root Mean Square Error
Private car	To	57.8	51.0	0.95	35.9	57.8	57.6	0.96	35.7
	From	45.8	39.2	0.93	34.6	45.8	45.0	0.95	29.3
Taxi	To	22.9	30.2	0.95	28.8	22.9	23.1	0.95	28.1
	From	17.6	35.2	0.97	42.5	17.6	17.4	0.97	17.0
Limousine	To	8.2	7.6	0.65	15.6	8.2	8.1	0.67	15.6
	From	14.5	3.4	0.80	27.8	14.5	14.4	0.87	18.2

TABLE 6

COEFFICIENTS FOR SELECTED MODELS

Mode	Direction	Business Trips			Nonbusiness Trips		
		Time	Cost	Constant	Time	Cost	Constant
Private car	To	-0.0401	-1.9097	0.0	-0.0342	-1.6058	0.0
	From	-0.0276	-1.3031	0.0	-0.0144	-0.6767	0.0
Rented car	To	-0.0265	-0.6030	1.7187	—	—	—
	From	-0.0150	-0.3505	-0.9498	—	—	—
Taxi	To	-0.0624	-0.3962	2.0495	-0.0631	-0.4025	0.3610
	From	-0.0642	-0.3447	1.2519	-0.0573	-0.3082	1.2998
Limousine	To	-0.0224	-1.0646	1.6664	-0.0177	-0.7922	-1.0416
	From	-0.0106	-0.5482	-0.1488	-0.0104	-0.5348	-0.0237

each mode, the coefficients in many cases did not meet the sign condition. The transformation of time into cost was also applied as in the case of the maximum likelihood calibration technique. This transformation did actually yield fewer incorrect coefficients, but the results were still unacceptable. It might be of interest to note that the private car coefficients (which were negative) remained at the same order of magnitude as those derived by maximum likelihood.

Testing

The model predictions were compared to the observed values of the sample. This comparison was performed both at the modal-split level and the trip level for each mode and each stratification. The latter comparison is, in fact, a weighted formulation of the former inasmuch as the larger the total trip number is for a given observation, the higher is the importance attached thereto. For each of the 4 calibrated models, data given in Tables 4 and 5 display for each mode the observed and estimated average trips, the correlation coefficient between estimated and observed trips, and the root mean square error of the prediction.

In view of the prototypical nature of this model, the relative importance of this task within the overall objectives of the study, and the nature and quality of the calibration data, it was decided to select the models described in the preceding because these models display desirable attributes in terms of sensitivity. The modal-mean trips were matched and the dispersions were reduced for each mode by adding an empirical constant to the linear forms attached to each mode. Because of the structure of the model, it is sufficient to find three such constants for the business models and two for the nonbusiness models. These constants are given in Table 6 and the results of their incorporation into the model are given in Tables 4 and 5 parallel to those of the unadjusted models.

CONCLUSIONS

The work discussed in this paper represents a first step toward the development and application of a modal-split model that is not limited as to the number of competing modes. Although developed for an airport access application, it is equally valid for travel mode analyses at the intraurban or intercity levels.

In contrast to other formulations, the elasticity of the share of a given mode is not constant but rather is proportional to the share of that mode and to the value of the independent variable with respect to which the elasticity is considered.

Furthermore, the underlying assumptions in the formulation of the model appear desirable because a mode is most "vulnerable" or most "attractive" when it holds

half of the market and because its "potential" of gaining (or losing) patronage is minimal when its share is already large (or small).

Other researchers have encountered difficulties in calibrating multimode modal-split models by relying on linear regression techniques. The present work suggests that use of maximum likelihood techniques can be helpful in alleviating these difficulties.

Experience derived from this study suggests that the model can be a useful tool for predicting modal split. It is felt, however, that additional data could improve the model's forecasting capabilities in an airport access application. In particular, as mentioned earlier, it appears important to know whether a traveler is a resident of the region under study. Several other measures also could be used. They include reliability of a mode as reflected by the standard deviation of travel times; income, which greatly determines the ability to pay; trip duration, which would give better information on parking cost; and group size, which is necessary for better cost estimation. Furthermore, smaller and more homogeneous zones could contribute to more accurate results.

Finally, consideration should be given to the dichotomy between the perceived and the observed values of impedances, particularly in terms of time and cost. This is all the more important in behavioral models inasmuch as the perception of impedances has necessarily some influence on user behavior and hence on model choice. A recent paper by Watson (5) gives an assessment of this problem and the implications on modal-split modeling resulting from biases in impedance data estimation.

In general, this initial investigation suggests that the n-dimensional logit model is a flexible and useful tool that can be applied operationally and that warrants further refinement. Despite the lack of certain key data elements, results of this airport access application were quite satisfactory. It is hoped that the results of further tests will confirm initial expectations.

ACKNOWLEDGMENTS

The work reported in this paper was performed by Peat, Marwick, Mitchell and Company in partial fulfillment of a study undertaken for the Metropolitan Washington Council of Governments and the U.S. Department of Transportation.

The authors would like to express their appreciation to several people. In particular, they would like to thank Peter Stopher for his considerable assistance in the early phases of the study. Our thanks go to Jeffrey M. Bruggeman and Robert L. November for their significant contributions to preparing the requisite computer programs and for the long night hours they spent in this effort. Finally, the authors would like to acknowledge the work of J. G. Cragg, who developed the initial version of the computer program for the maximum likelihood calibration.

REFERENCES

1. Finney, D. J. Probit Analysis. Cambridge Univ. Press, 1964.
2. Lisco, T. E. The Value of Commuter's Travel Time: A Study in Urban Transportation. Univ. of Chicago, PhD dissertation, 1967.
3. Stopher, P. R. A Probability Model of Travel Mode Choice for the Work Journey. Highway Research Record 283, 1969, pp. 57-65.
4. Reichman, S., and Stopher, P. R. Disaggregate Stochastic Models of Travel-Mode Choice. Paper presented at the HRB 50th Annual Meeting and included in this RECORD.
5. Watson, P. L. Problems Associated With Time and Cost Data Used in Travel Choice Modeling and Valuation of Time. Paper presented at the HRB 50th Annual Meeting and included in this RECORD.

PROBLEMS ASSOCIATED WITH TIME AND COST DATA USED IN TRAVEL CHOICE MODELING AND VALUATION OF TIME

Peter L. Watson, The Transportation Center, Northwestern University

The aim of this paper is to clarify some of the issues raised by the current disagreement as to whether it is more appropriate to use "perceived" or "measured" data in models that estimate modal choice. The procedure adopted is to consider the hypotheses that can underlie the modal-choice models and to discuss which data type is more appropriate, given the suggested hypotheses. From the basic premise that the data used to estimate the model should correspond to the variables specified in the hypothesis, 2 hypotheses are derived. The former relates choice of travel mode to the times and costs of the journey as perceived by the traveler, with the implication that perceived data should be used to calibrate the model; the latter is a special case of the former that relates choice to the expected times and costs of the journey, and it is suggested that the best available approximation to the expected values is the measured data. Problems arise when it is difficult to decide which hypothesis is more appropriate in a given case or when the group of travelers is heterogeneous with respect to the hypothesis that underlies their choice behavior. The problem resolves to one of the effects of using the wrong data set. In the case of predictive models, it is suggested that the choice of data is relatively unimportant because models based on perceived data are not recommended for predictive purposes, given the absence of information on the stability of the relationship between perceived and measured values and in the light of the problems of predicting perceptions. In the case of models used for the derivation of a value of time, the problem is more acute. A preliminary investigation indicates that the use of each data set produces a value of time that is biased with respect to the true value. It is argued that such a bias is not a problem provided that 2 conditions are fulfilled: (a) The data used are consistent with the behavioral hypothesis; and (b) care is exercised in the use of the value of time so that a value derived in 1 situation is not used to value time saved in a different situation. In conclusion, the answer to the question as to which data type is correct lies not in dogma but in the careful use of logic in the selection of the hypothesis and the use of the information derived from the model.

•THE PROBLEM considered in this paper arises from the use of different types of data to calibrate modal-choice models. To understand the problem, it is necessary to consider some aspects of the development of modal-choice models. In essence, the aim of such a model is to explain an individual's choice of travel mode (or route) in terms of a series of variables that describe the characteristics of the trip, the alternative modes, and the traveler and his household. Because the analyses are identical, the exposition will proceed in terms of modal choices.

Sponsored by Task Force on Value of Travel Time.

For the purposes of this exposition, it is unnecessary to describe in detail the statistical techniques that have been used in modal-choice models. Suffice it to say that the most commonly used techniques have been discriminant, logit, and probit analyses. The basic procedure is to relate a dependent variable, representing the choice made, to a number of independent variables as described in the preceding. The problem then arises as to which explanatory variables should be used. As a basic research technique, the analyst selects his explanatory variables on the basis of received theory or postulates about the relationship or does both of them, combines them in the light of previous work and his own experience, and then tests the model to investigate the explanatory power displayed by his selection. In the case of modal-choice models, the most commonly selected variables have been the times and costs of the journey by each of the alternative modes. (Because these variables pose the largest problems in this area, the analysis will assume that they are the only variables under consideration. It should be kept in mind that the discussion may apply to other variables.) The task of the researcher is then to collect data on times and costs by each mode for a sample of individuals to test the model. The problem arises because it quickly becomes clear that the time or cost of a journey is not a unique value. If the analyst measures the time for a given journey, he obtains 1 value. If he asks the traveler how long the same journey has taken, he might obtain a different value. The same findings emerge when data on costs are collected. It is evident that the individual traveler does not always know exactly how long a journey has taken or how much it has cost, and this fact leads to the dichotomy between perceived values and measured data. It should be noted at this point that much confusion has been caused by the use of multiplicity of terms to refer to these 2 values. For the purposes of this paper, the terms "perceived" and "measured" will be used where "perceived" refers to the time or cost as perceived by the traveler (such values have variously been called "behavioral" or "subjective" data) and where "measured" refers to the value obtained by the analyst replicating the journey either physically or by means of printed matter such as timetables or fare schedules (such values have been called objective, real, engineering or actual values). It is intended to further distinguish the true times and costs that may differ from both the perceived and the measured times and costs.

The following discussion considers the problems that may arise in the collection of perceived and measured data that arise in the collection of such data or that may be inherent in the data. The analysis that follows these findings will consider whether perceived or measured values are more appropriate for inclusion in a modal-choice model. It will also consider the effects of the biases on the modal-choice modeling process and on the derivation of a value of time.

BIASES IN THE DATA

The word bias is perhaps a misnomer in the sense that it implies a degree of undesirability that is not intended. The term bias is used to refer simply to the differences between the various measures and estimates of times and costs; for example, the true times and costs of a journey and the traveler's perception of those times and costs. There is ample evidence (2, 3) to demonstrate that such differences do occur. The problem is to identify their magnitude, their direction, and any factors that might cause them. Initially, only biases in the perception process will be considered. At a later point the possibility will be introduced that further biases may appear when the perceived value is reported and when objective measurements are made. The postulated bias may be different in different circumstances: notably whether the times and costs refer to the chosen or the rejected mode, or whether the traveler is familiar with the mode. The following discussion will consider several types of error in turn and will contain comments on the ways in which each might be affected by changes in circumstances.

Perception Bias

The bias to be considered in this section is the cornerstone of the errors generated in the perception process. Other biases may aggravate the situation, but it is suggested

that, at best, they only modify the perception bias. The typical traveler is not a work-study engineer; i.e., he does not travel with a briefcase in 1 hand and a stopwatch in the other. As a result, he often does not know how long a certain journey has taken because it is notoriously difficult to estimate elapsed time accurately. Therefore, his perception of the time taken on a journey may differ from the true time taken. As a best approximation, it may be assumed that such errors will be random, be approximately normally distributed, and have an expected value of 0. There is some evidence to suggest that this may be so (3), but it is by no means an established fact.

It may be thought that the problem might be less acute when perceptions of costs are concerned because actual payments are involved, and there is some evidence to support this view (2). The conclusion is somewhat obscured by the problem of allocating costs to different journeys, both when the car is involved and when some type of season ticket is used for transit movements. Little can be said about the biases resulting from such allocative problems. Lansing and Hendricks (2) found that people consistently overestimated the running costs of their cars, a result that conflicts with the often-postulated view that such costs are always underestimated. It is possible that, like time biases, cost biases can be assumed to be random, normally distributed with 0 mean.

It is, of course, possible that the perception bias could be affected by the learning process. One could postulate that the bias is a decreasing function of the traveler's familiarity with the modes. This would mean that the perception bias would be smaller for the greater degree of the traveler's familiarity with the mode. In other words, the traveler may misperceive the time taken to drive to work, but his perception bias is likely to be greater for a journey that he has undertaken only a few times, and even greater for a journey that he has never undertaken. In the limit, it would approach a guess based on little or no information about the journey, e.g., if the journey involved traveling to an unfamiliar destination over unfamiliar roads, when even the distance to be traveled may be unknown.

Such a postulate would not necessarily imply that the relationship varied as the mode was chosen or rejected. For the work trip, it seems likely that the traveler would be familiar with his chosen mode, but, for an infrequent trip, he may be unfamiliar with all the possible modes. It seems not unlikely, however, that, even for a frequent trip, the traveler may be ignorant of the characteristics of the mode that he does not habitually use.

Rounding Bias

The rounding bias arises from the fact that people appear to find difficulty in distinguishing small units of time, like a minute. This leads them to perceive time in lumps of, say, 5 min. It is clear that a motorist, for example, does not think in terms of an 18-min journey, but in terms of one of 15 or 20 min. Thus, there is an observable tendency to round journey times up (or down) to multiples of 5 min. It is possible that on longer journeys, times are rounded to 10 min, 30 min, 1 hour, or even larger units.

Whereas it is quite clear that rounding biases exist for journey times, it is less clear that they exist for costs. However, it seems possible to assume an analogous process in which people round their costs to multiples of, say, 5 cents. In the Edinburgh-Glasgow area modal-split study, that is currently being conducted under the auspices of the British Ministry of Transport, this was not observed, and costs were typically not rounded; but the phenomenon may be observed for longer trips.

Given the postulate that five is the rounding unit, it would be simple to predict the rounding bias if the direction of bias were known. It has been suggested that people are more likely to round down for the chosen mode and up for the rejected mode, but no substantial evidence exists on the subject. Because discussion of this possibility leads into consideration of reporting problems, it is intended to deal with it at a later point. In the present context, it is sufficient to say that rounding as part of the perception process seems as likely to follow a simple rule (e.g., round to the nearest multiple of five) as it is likely to follow a more complex one.

Antimodal Bias

It is possible to postulate that the traveler's perceptions of times and costs for the rejected mode may be subject to a bias that results from the traveler's feelings about the mode. Should the traveler dislike the rejected mode intensely, his perception of the times and costs by that mode may be subconsciously influenced by that dislike. If he is unfamiliar with the mode, the unfamiliarity, combined with the adverse feelings, could account for the large errors that have been observed to appear when times and costs by rejected modes are reported.

Such a bias could be referred to as an attitudinal bias, based as it is on the traveler's attitude toward a given mode of transport. It should be emphasized that little or no work has been done on this subject, and therefore, little can be said about the nature of the bias. This may, however, be an important feature of the perception process.

Reporting Bias

The discussion so far has proceeded on the assumption that the perceived times and costs are the numbers collected by interviewing travelers. It is now appropriate to relax that assumption. It is postulated by Thomas (5) that a bias appears in the reporting of times and costs. In other words, the times and costs of a journey as perceived by the traveler are modified when he reports them either to an interviewer or in a questionnaire. The reporting bias arises from the traveler's desire, conscious or (more probably) subconscious, to make his actions (i.e., his choice) appear more realistic, more reasonable, or more impressive or to gain the interviewer's approval for his actions. Given a trade-off situation, it is argued, he will attempt to make either the time saving for a given expenditure seem greater or the expenditure for a given time saving seem smaller or both. These objectives can be achieved in a number of ways that reduce to various combinations of underreporting times or costs by the chosen mode or both, or the overreporting times or costs by the rejected mode for both.

Thus, the postulate that this bias exists implies the direction in which it will operate. It is, however, not possible to say anything about the magnitude of a bias because the perceived times and costs are not known.

Normalizing Bias

It is possible that, for a given journey, the traveler may try to correct for atypical journey times; e.g., he may report the time normally taken for the journey. (Clearly, this problem can only arise when the subject is familiar with the mode and the journey.) It is felt that careful designing of the questionnaire and training of the survey staff can eliminate this problem.

In the previous sections, a number of possibilities have been suggested that might explain discrepancies between the true times and costs and the times and costs as perceived and reported by the traveler; these biases could be interactive or cumulative. However, it cannot be too strongly stressed that little work has been done on this topic, and, therefore, the postulated biases are based on a mixture of propositions and casual empiricism that may not be demonstrable. The strong evidence that exists (4, 5) shows only that the data, as reported by the travelers, differ from the data as measured by the analyst. It is possible that both sets of data differ from the true value and it is, therefore, appropriate to consider the ways in which the measured values may differ from the true values.

Measurement Bias

It may be assumed that biases in measurement derive from the manner in which the measurements are made and not from the incompetence of the person taking measurements. This being the case, it follows that, if the measurer accompanied the traveler on his trip and measured all the time and money expenditures, then the measured

times and costs would correspond exactly with the true times and costs. In practice, such a procedure is not practicable and the so-called measured times and costs are usually estimated from standardized data, based on fifth-wheel test runs, average speeds, assumed routes, fare schedules, and published automobile operating costs. In other words, the times and costs are derived on the basis of a series of conventions. For example, the time spent walking to a station may be based on the distance and an average walking speed; the time spent driving a given route may be based on the distance and average driving speeds; the cost of a trip may be based on the distance and an average mileage cost for the given vehicle; and the time spent waiting for a bus may be based on half the scheduled headway or 10 min, whichever is greater. The result is that the measured time corresponds not to the actual time taken by the traveler but to the time that he would take on an average for the given journey.

CHOICE OF THE CORRECT DATA BASE

Given that some combination of the biases outlined previously leads to a difference between perceived and measured times and costs, the question arises as to whether perceived or measured data are more appropriate for use in the testing of a modal-choice model. Considerable controversy has arisen over this point, with the combatants polarizing into 2 camps that may be characterized in the following way. The advocates of the use of perceived data argue that the modeling procedure attempts to delineate a statistical relationship between the choice of an individual and the factors influencing that choice. Given such a starting point, it is logical to argue further that, if factors such as times and costs are to be included, the appropriate times and costs to use are those perceived by the traveler. In such a situation, it is illogical to use measured data because the traveler does not know what the measured values are, and the use of such data implies that he bases his choice decisions on data that he does not possess a conclusion that is curious, to say the least.

The advocate of the use of measured data, on the other hand, would argue that the precise process by which the modal-choice decision is arrived at is irrelevant and that the important feature of his model is that people behave "as if they based their choices on measured data." Moreover, a model based on measured data can predict what people do with a high degree of accuracy. Such an argument, combined with the reduced problems of data collection associated with measured data, constitutes the case for its use.

It would be easy at this point to dismiss the problem as an insoluble controversy between the theorist on the one hand, who wishes to explain or replicate the decision-making process and the pragmatist on the other hand, who simply wishes to predict or replicate observable behavior. Such a solution to the problem is, however, unsatisfactory in the sense that prediction is a more satisfactory procedure if the model to be used for prediction is based on a sound hypothesis. The fact that a model predicts may well be a statistical accident, and the absence of a sound hypothesis may inhibit the analyst's confidence regarding the stability of the relationship and, hence, his confidence in the predictions. It would clearly be more satisfactory if the use of measured data could be justified on the basis of a behavioral hypothesis, similar to the one that underlies the use of perceived data.

Such a hypothesis is now suggested. The basic hypothesis that the individual bases his modal-choice decision on times and costs as he perceives them is maintained, but a special case of perception is defined. In the situation where a trip is undertaken regularly, it is argued that a learning process will operate such that the times and costs as perceived by the traveler can be represented as the times and costs he expects to be faced with on the given trip. This is to imply not that decisions will be based on "expected values" in a rigorous statistical sense (with all experiences being given equal weight) but simply that the traveler, as a result of frequent travel, comes to expect that a journey will take a certain time and cost a certain amount. Moreover, the discussion on measurement biases concluded that the measured times, for example, correspond to the time that the traveler would take on the average for the given journey. Thus, a close relationship is demonstrated between measured data and perceptions in

the special case, where perceptions can be interpreted as expected values based on a learning process and repeated exposure to a given trip. It is clear that the special case of hypothesis lacks some of the generality of the perception hypothesis, because the fact that the expected values are based on experience implies that the traveler has some degree of familiarity with the trip characteristics.

Having given the use of measured data a certain degree of respectability in the form of its own hypothesis, we should appropriately return to the question of which data base is correct. It is argued that the use of a given set of data is only correct when it represents the variables included in the model. Given such a test of correctness, the use of measured data in a model that represents infrequent trips is clearly incorrect. However, the converse is not true. Because the hypothesis supporting the use of measured data is a special case of the perception hypothesis, either of these types of data may be held to be correct in a model representing a familiar journey, and the choice may be based on other factors that will also be of importance in choosing the data types when the population does not fall into a familiar or unfamiliar category but contains some of each type. These factors may be pragmatic, such as ease of collection or the avoidance of reporting biases, or, preferably, may be based on a consideration of the implications of using each of these types of data for constructing modal-choice models and for the values of time that may be derived from such models. These implications are considered in the following sections.

IMPLICATIONS FOR MODAL CHOICE

Perhaps the most important use of a modal-choice model is in prediction. A model can be used to predict what will happen to the modal split in a given situation if the trip characteristics are modified in terms of either the costs or the times of the journey. The policy-maker may wish to know what will happen to the relative numbers of people using each mode if, for example, rapid transit is made faster or cheaper, or if, on the other hand, traffic restrictions make automobile travel slower. In such situations, data corresponding to the new state can be fed into the model to estimate the modal-split changes.

In the same way, data from a different location could be used to estimate the modal split given the characteristics of the new location. Thus, the model can be used to analyze data that are different either temporally or spatially from the data on which the model was based.

For such a procedure to be effective, certain conditions must be met. There must be some degree of confidence that the relationship represented by the model remains constant over time and space. If there is reason to believe that the relationship may change, then the procedure is meaningless. Furthermore, the data from the new situation or location must be of the same type as the data used to calibrate the model in the first place. Given the fact that it is frequently difficult to obtain data, which reflects precisely the variables sought to be included in the model, the success of the estimating procedure may be partially dependent on the precise nature of the data. For example, many economic relationships depend on an income variable, and, in many cases, it is not often possible to measure, say, disposable income precisely. Therefore, an estimate must be used, and the relationship may be established by using 1 estimate and shown to be spurious by using another. In the same way, a modal-split model calibrated with perceived data might be dependent on the nature of the variable. A change to a new location or situation could change the nature of the variable because it is unknown whether the difference between perception and reality (perception bias) is constant over time or space. To the extent that it is not constant, the analyst could have less confidence in his predictions.

These problems seem likely to be most acute when analysis of new modes or major service changes is undertaken. It is difficult, if not impossible, to envisage a method by which the public's perception of the reality of a new situation could be predicted. Thus, the problem reduces to a very simple one: To predict future values of the dependent variable in a relationship, one must first know the values of the independent variables. It is, given present knowledge, not possible to predict future misperceptions.

It should be noted at this point that the use of measured data obviates this problem because, if the model were calibrated with measured data, there would be no reason to believe that the nature of measurement biases would vary from place to place or from time to time. Moreover, the problem would be greatly simplified if a stable relationship between perceived and measured data could be found.

IMPLICATIONS FOR THE VALUATION OF TIME

An important spin-off from the modal-choice model is a value of time, which is often used to evaluate time savings in cost-benefit types of analyses of transportation investments. Because time savings frequently represent a high proportion of total benefits, the total value of benefits, and, hence, acceptance or rejection of a project is very susceptible to small changes in the value of time. This section analyzes the effects of using the different data bases to estimate the modal-choice models from which values of time are derived.

To analyze the effects on the value of time of using different types of data requires a consideration of the mechanism by which the value of time is derived. The discussion can be simplified by considering the simplest type of model, one that represents a linear relationship between the probability of a given choice and the times and costs of the alternatives expressed as differences:

$$P(x) = \alpha_1 \Delta T + \alpha_2 \Delta C$$

The value of time is then defined as the rate at which time would be traded for cost to leave P(x) unchanged, i.e., α_1/α_2. It is contended that, because the relationship is based on the observable behavior of the individual travelers, the trade-off ratio represents the rate at which the population as a whole, aggregated by the chosen statistical procedure, will trade time for cost, i.e., the value that is placed on time.

The question that then arises concerns the effects of using different types of data on the value of time thus derived—a question that reduces to a problem of assessing the effects of the different data types of the coefficients estimated by the model. The following example, which is a modified and extended form of one used by Thomas (5), shows some of the effects that might be expected.

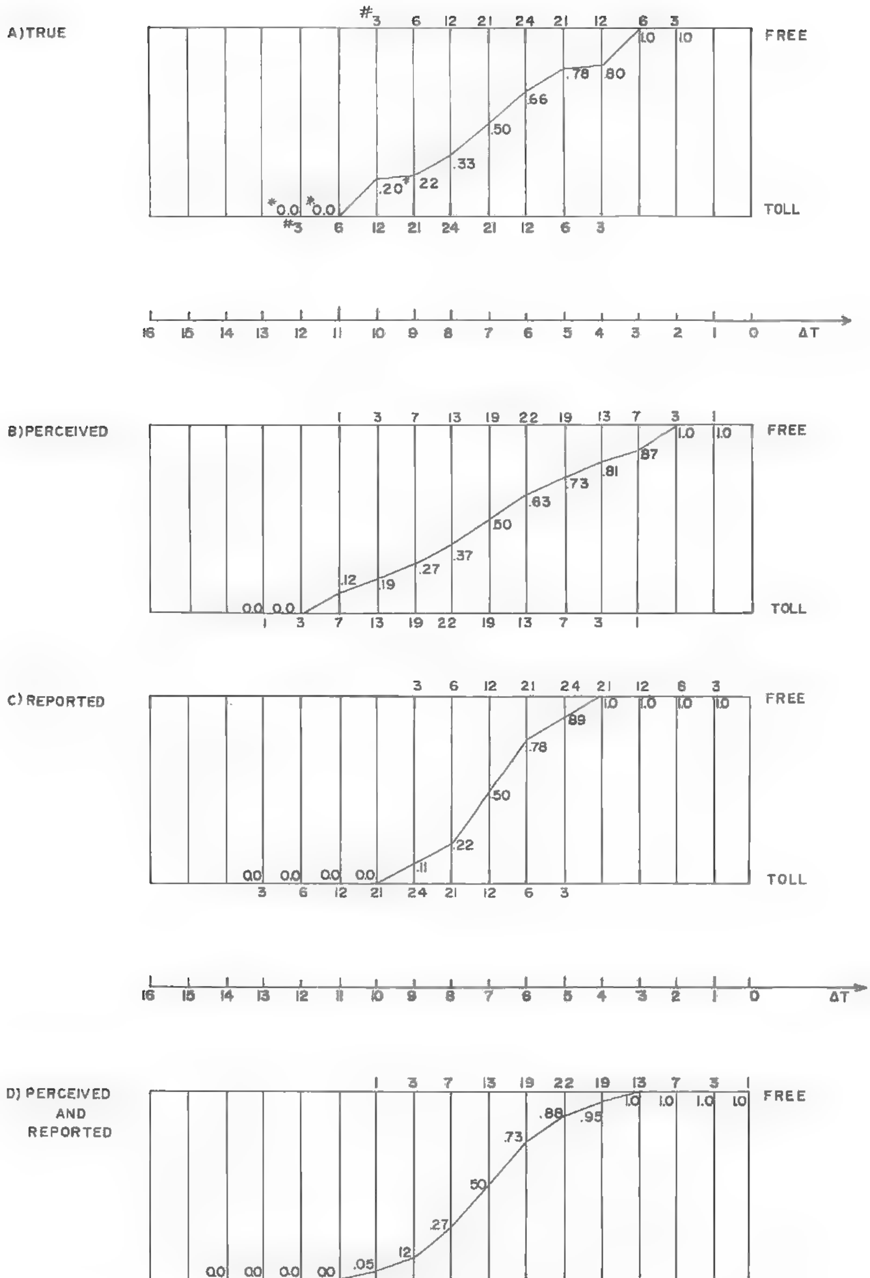
The exposition can be clarified by considering in isolation the effects on the coefficient related to the time difference variable because such a procedure greatly simplifies the problem and permits a graphical treatment in 2 dimensions. The situation envisaged is one in which a number of motorists have faced a choice between a free road and a 10-cent toll road. Faced with a difference in travel time between the 2 roads, all other things equal, a choice has been made. Figure 1a shows the observed modal split for each time difference and the curve that is estimated by using the true time differences. It should be noted that this curve is a theoretical construct to be used only as a yardstick; because the true values are unknown, it is not possible in practice to estimate this curve. This curve represents a part of the choice model, namely,

$$P(x) = \alpha_1 \Delta T$$

and thus the slope of the curve, which represents the rate at which P(x) changes for a unit change in ΔT, is equivalent to α_1 or to the numerator in the expression for the value of time:

$$VOT = \frac{\alpha_1}{\alpha_2}$$

For the purposes of this example, it is assumed that, if the cost difference (or toll) is held constant, the coefficient corresponding to $\Delta C(\alpha_2)$ will remain constant; and thus changes in α_2 due to the use of different data types will involve changes in the value of



#Numbers observed on free and toll roads at given time differences (ΔT).
*Proportion choosing free road.

Figure 1. Effects of biases on a_1.

time. Moreover, because the slope of the curve relating P(x) to ΔT is equivalent to α_1, a change in the slope of the curve will change the value of time, a steeper slope implying a higher α_1 and, hence, a higher value of time. It is recognized that this assumption will not hold because only the true value of α_2 will remain unchanged whereas its estimate, $\hat{\alpha}_2$, may vary. It is argued that, even given the heroic nature of the assumption, the analysis will indicate the directions of changes in the value of time and the magnitude of the resultant problems.

Figure 1 shows the effects on the slopes of the curves developed by using different data types. Figure 1b shows the effects of misperceptions of the "true" times. In the absence of any better information about such biases, it was assumed that, for each time difference, one-third of the travelers would perceive the time difference as 1 min greater than it really was, one-third would perceive it as 1 min less, and one-third would perceive it correctly. (It is not claimed that this representation of bias has behavioral validity; it is used merely as an example.) It will be clear that the effect of these hypothetical biases is to render the slope of the curve less steep and thus to reduce the value of time derived from such data compared with the value derived from the true data.

Figure 1c shows the effect of a reporting bias, as a result of which each traveler misreports the true time difference by 1 min in favor of his chosen mode. Thus, the toll-road chooser makes it appear that he gained more time as a result of paying the toll, and the free-road chooser makes it appear that the time saving he gave up was smaller than it really was. The effect of such a systematic bias is to make the slope of the curve steeper and, thus, the value of time higher.

Finally, Figure 1d shows the effect of combining the perception and reporting biases. The resultant slope is less steep than the reported slope but steeper than the perceived slope. It is clear that the steepness of the slope will be dependent on the degree to which the resultant biases offset each other, because they work in different directions.

The use of notional perception and reporting biases in the example reflects the author's view (based on little more than casual empiricism) that these biases are likely to be of greater importance than the other biases outlined in the section on possible biases. Should the reader disagree, the mechanism for analysis is still valid and different propositions about the nature of the biases can be tested.

After this consideration of the biases that might occur in data collected from the travelers themselves, it is now possible to proceed to a consideration of the biases that might occur in the measurement process. This was done by following the same procedure. A number of assumptions were made about the ways in which the measured values might differ from the true values. Three assumptions were made:

1. It is assumed that the necessity of duplicating test runs will lead to some level of aggregation of travelers, with a resultant reduction in the time differences at which travelers are observed. The aggregation is arbitrary in that all the odd-numbered time differences are reallocated equally to the even differences above and below them.
2. Lest the experiment be influenced by the nature of the aggregation, No. 2 aggregates to the odd-numbered time differences.
3. In this measure, the reasoning used by Thomas is followed, and it is assumed that the mean time difference of each group remains unchanged while the variance of each group increases.

It can be shown in all 3 cases that the slope is observed to be less steep than the slope based on the true data, reflecting that the value of time is less than the value that would be derived from the true data.

It should be remembered that, so far, the problem has been very much simplified, in that we have assumed a route-choice problem with a fixed cost difference, ΔC, and no difference in operating costs between the routes. If the operating cost assumption is relaxed, or if the analysis is extended to a modal-choice situation in which the traveler may well be faced with costs as well as times that vary from mode to mode, the problem becomes very much more complex. It does not seem unreasonable to expect that the biases appearing in the time data should also appear in the cost data.

If this is so, it is not only the numerator of the value of time expression

$$\text{VOT} = \frac{\alpha_1}{\alpha_2}$$

but also the denominator that may vary. The implications for the value of time derived from a model under such circumstances are very serious because the interactions of the various types of time and cost biases lead to complex effects on the value of time. The following data give the direction of the effect on the value of time of different combinations of time and cost biases. It should be noted that in some cases not even the direction of change can be predicted because it is dependent on the relative magnitudes of the biases.

Time (α_1)	Cost (α_2)	VOT (α_1/α_2)
Increase	No bias	Increase
Decrease	No bias	Decrease
No bias	Increase	Decrease
No bias	Decrease	Increase
Increase	Increase	?
Decrease	Decrease	?
Increase	Decrease	Increase
Decrease	Increase	Decrease

It should be stressed that, if the problem of cost biases appears to have been given a rather cursory treatment, it is not because it is felt that such biases are unimportant but rather because the arguments put forward in the discussion of time biases apply equally to cost biases, and it would be tiresome to repeat them. It should be noted, however, that the problem of time biases, complex as it may be, is but a part of a problem whose complexity is greatly increased by the introduction of cost biases. This increase in complexity is due to the multiplicity of interactions and cumulations that are possible among the different biases.

Is it, then, possible to select 1 value of time as the correct one? Indeed, it is furthermore necessary to decide what is to be meant by "correct." It is contended that a value of time based on perceived data (different though it may be from one derived from true data) is correct if it is derived from a model whose hypothesis is such that it relates choice to the values of times and costs as perceived by the traveler. In the same way, a value of time based on measured data is correct when it is derived from a model whose hypothesis is of the special case that relates choice to an expectation of times and costs based on familiarity, which can be represented by measured data. Thus, 2 values of time may both be correct in the sense that they are both derived from properly constructed models.

CONCLUSIONS

A number of interesting conclusions can be drawn from the preceding discussion:

1. The development of the special case hypothesis means that the use of both measured and perceived data can be considered correct for use in modal-choice models in the sense that a meaningful hypothesis about traveler's behavior can be suggested to justify the use of each type of data. Because the appropriate hypothesis can be selected on the basis of traveler's familiarity with the trip under consideration, the selection of the data types poses no problems in cases where a traveler is familiar with the trip. In other cases, 3 factors are of importance: the stability of the relationship between measured and perceived values and the effects of using the inappropriate data types on predictions and on values of time. In terms of the use of the model to predict travel behavior, the problem is less acute as a result of the conclusion that the use of perceived data can lead to serious problems in the prediction process (i.e., that instability

in perception errors over time or space could lead to a lack of confidence in the predictions).

2. If it can be demonstrated that the relationship between measured and perceived data is stable, a procedure might be developed for modifying measured data into perceived data for those cases for which the use of perceived data was indicated. Should a procedure prove impossible to establish, or if it is impracticable to operate, it is concluded that an investigation of the effects of perception errors on predictions must be undertaken before models based on perceived data can be used with confidence for prediction purposes.

3. In terms of the use of the model for the provision of values of time, it is felt that models based on both perceived and measured data can produce values of time that are correct in the sense that they are derived from properly constructed models. It is suggested that, in the same way as values of time are beginning to be classified by income group and trip purpose, they be further classified by familiarity with the choice situation. (It is, of course, possible that this variable is linked with trip purpose.) In this way, each value would only be used in a situation similar to the one in which it was derived. It is argued that, just as it is wrong to apply the value of time derived from a commuting trip to time saved on a pleasure trip, or to apply a value of time saved derived from the choices of upper income group travelers to time saved by lower income group travelers, it is equally wrong to apply a value derived from a familiar situation to time saved in an unfamiliar situation. The conclusion that values of time based on both perceived and measured data can be regarded as correct represents the addition of a further dimension to the growing matrix of values of time.

4. The answer to the question as to which data types are correct cannot be answered by dogma but must be answered by analysis of the choice situation under consideration and a decision as to which hypothesis about traveler decision-making behavior (and, hence, which data types) is more appropriate in a given situation.

REFERENCES

1. Harrison, A. J., and Quarmby, D. A. The Value of Time in Transport Planning: A Review. European Conf. of Ministers of Transport, Round Table, 1969.
2. Lansing, J. B., and Hendricks, G. Living Patterns and Attitudes in the Detroit Region. Detroit Region TALUS Report, 1967.
3. Lisco, T. E. The Value of Commuter's Travel Time. Univ. of Chicago, unpublished PhD dissertation, 1967.
4. Quarmby, D. A. Choice of Travel Mode on the Journey to Work. Jour. of Transport Economics and Policy, Sept. 1967.
5. Thomas, T. C. The Value of Time for Passenger Cars. Stanford Research Institute, Menlo Park, Calif., 1967.
6. Thomas, T. C., and Thompson, G. I. The Value of Time for Commuting Motorists as a Function of Their Income Level and Amount of Time Saved. Highway Research Record 314, 1970, pp. 1-19.
7. Watson, P. L. The Choice of the Data Base for Modal Split Modelling and the Valuation of Time. Planning and Transport Research and Computation Company, Urban Traffic Model Research Symposium, London, 1970.

ANALYSIS OF TRAVEL PEAKING

William Ockert and Richard Easler, Regional Planning Council, Baltimore, Maryland; and
Franklin L. Spielberg, Alan M. Voorhees and Associates, Inc.

Transportation planning simulation models have generally been structured for total daily travel. A question that has seldom been raised is the extent to which peak travel patterns differ from 24-hour patterns. This paper describes a modeling technique that can help in evaluating future transportation plans and programs by directly simulating peak-period demand. The model basically converts 24-hour trips to peak-period trips then allocates them to modes, and assigns them to networks. Analysis of the model's output indicates that the orientation of peak-period travel is significantly different from that of 24-hour travel. Nine communities within the urbanized area of the Baltimore region were chosen to illustrate peaking characteristics for different types of movements, i.e., peaking from employment areas or peaking from residential areas. Comparison showed that projected peaking changes based on base year observation cannot provide reasonably accurate estimates of future peak-period conditions. Peaking characteristics were found to change through time as a result of the uneven growth in employment and population. Based on the analysis, it is clear that peaking factors change through time and are sensitive to the distribution of urban activities. Thus, the use of the peak-period simulation model will eliminate numerous errors in estimating future travel conditions.

•THE RUSH-HOUR PROBLEM—congestion, crowding, delays, and substandard speeds—not only creates frustration and tension for travelers but also leads to economic stagnation of the urban area and thus aggravates its social problems. In the Baltimore region, nearly 40 percent of all travel occurs in only 4 hours of the day, those hours during which people must travel to or from their jobs. In fact, 60 percent of all work trips in Baltimore occur in this period (according to data derived from a 1962 origin and destination study). Thus, people making the most repetitious trips and the ones least subject to personal adjustment face the worst travel conditions.

In recent years, transportation planners and urban area system designers have focused on the development of highway and transit facilities that will satisfy long-term demand within the constraints imposed by other elements of the urban system—resources and environment, for example. The projection of demand has generally been developed by using a series of generation, distribution, modal-choice, and assignment models, the techniques of which are well known.

These models, however, have generally been structured for total daily travel, although it is recognized that the requirements for most facilities are set by the peak-period demands. Many methods have been devised to bridge the gap between 24-hour and peak-period travel. Some of these have dealt with traffic on a particular route; others have considered the entire region; some have merely extrapolated present trends. A question that has seldom been raised is, to what extent do peak-travel patterns differ from 24-hour patterns. Shifts in patterns could lead to underdesign of some facilities and excess, unused capacity for others.

Sponsored by Committee on Transportation Forecasting and presented at the 50th Annual Meeting.

For transit system design, in particular, peak-period patterns are critical because this period constitutes the major transit travel market. In the peak, when the majority of work trips are made and when highways are most congested, transit is most competitive. Failure to recognize peak-demand patterns can lead to incorrect investment decisions.

Different land use and transportation alternatives result because of variations in the amount and location of peak-period travel. In evaluating alternatives, peak-period travel conditions encountered by various segments of the population should be taken into account. For example, concentrations of activities that generate highly peaked travel are much more likely to cause severe traffic problems than activities of the same density that produce travel spread more evenly throughout the day.

PRIOR ESTIMATING METHODS

Transportation planners are familiar with the basic methods of estimating peak demand on specific facilities. Thirtieth highest hour distributions have been in use for years as a design standard. More refined estimates using peak-hour factors, K, and directional distribution factors, D, by facility also represent a popular method that has been the focus of many studies.

These methods, although applicable to short-term design, are difficult to relate to overall travel patterns and may fail to consider the changing nature of the region. A more refined method, based on the relationship between peak-hour factors, K and D, and the percentage of work trips on each facility, has been developed (1). This method approaches an analysis of trip patterns but is still focused on individual links. Others have applied area-wide factors, or factors for subareas, to either total trips or highway and transit trips separately. These methods tend to be unresponsive to changing conditions.

The best approach, perhaps, is one that involves a separate peak-travel model set. This approach deserves more research. In practice, it has been difficult to develop stable trip generation and distribution relationships, perhaps because the proper data, such as the relationship between peak-period travel times and trip-distribution patterns, have not been available. Furthermore, this approach requires a second complete model for total daily travel, a costly luxury for many studies.

The peak-period model described in this paper does not eliminate all difficulties either in logic or in data needs. It presents an approach to the problem that appears to give reasonable results both in base-year application and in future-year projection. It is presented to acquaint others with a different technique and to indicate the factors that, in this study, were found to be significant.

ANALYSIS AND MODEL FORMULATION

In the development of a set of transportation planning and evaluation models for the Baltimore Region, it was necessary to devise a method for estimating peak-period travel that is both responsive to changing conditions over time and also reflects the different peaking levels on individual facilities in the region. The methodology selected was to directly simulate peak-period demand. The developed model deals with overall person trip patterns and examines peaking of travel by trip purpose regardless of travel mode. The model basically converts 24-hour person trip tables to peak-period person trip tables, which are then allocated to modes and assigned.

The requirement that the temporal as well as the spatial distribution be considered raises many difficult questions. It is obvious that peak travel occurs at different times in various sections of the region; that, over the peak time span, the "demand" is limited by the capacity; that factors, such as working hours that are beyond the scope of transportation planners, will influence peaking conditions; and that, for certain modes (transit in particular), the peak travel time is set in part by scheduling practices.

It was felt, however, that, for long-range planning purposes, it was neither necessary nor within the scope of data reliability to attempt to pinpoint the peak over a short duration. Rather, the peak travel over a 2-hour period was deemed sufficient for demand analysis and evaluation. Although peaks on individual facilities within this period will

exceed the average 2-hour demand, they are likely to be of short duration and subject to variation with normal demand fluctuation. They are best treated by recognizing and allowing for the inherent errors in projection methodology and data.

The basic travel data for the model development were obtained in the 1962 origin and destination study conducted for the Baltimore Metropolitan Area Transportation Study (BMATS). These travel data were supplemented by social and economic data developed by the Regional Planning Council.

Analysis of the travel data revealed that, for total person trips, the peak 2-hour period of the day occurred between 3:30 and 5:30 p.m. (Fig. 1). This period includes not only the majority of trips from work to home but also many school trips, shopping trips, and trips for other purposes not found in the morning peak.

WORK TRAVEL ANALYSIS

Variables Examined

The strategy for model development involved a stratification of trips into 4 purpose categories—work, school, and other home-based, and other non-home-based. After this, an analysis was made of the degree of peaking of each purpose as related to individual variables. Among the variables examined in this stage of the investigation were trip-end variables such as production zone income, production zone residential density, attraction zone employment density, attraction zone employment composition; interchange variables such as trip time and distance; and trip-maker variables such as occupation and industry.

A comparison of the percentage of work travel in both the morning and afternoon peak periods versus the residential density in the production zone clearly indicated that there is only a slight change in peaking within density ranges.

Similar relationships were observed for other individual trip-end and trip-interchange variables. The individual variable that appeared to have the most influence on peak period travel was the industry in which the trip-maker is employed (Fig. 2).

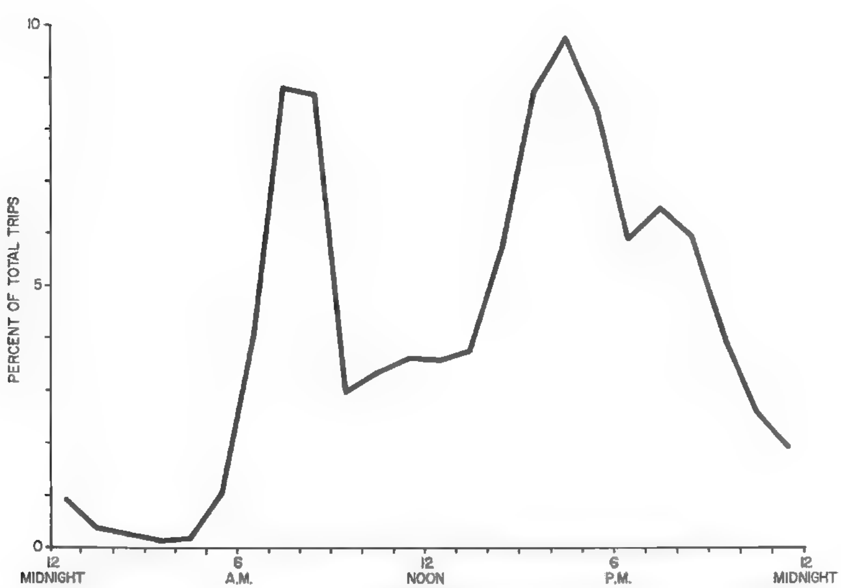


Figure 1. Person trips by reported hour beginning.

Development of Relationship

Logically, those persons employed in industries (government, wholesale, manufacturing, and construction), which tend to maintain standard working hours beginning between 7 and 9 a.m. and ending between 4 and 6 p.m., have the greatest tendency to make their work trips in the peak period. On the average, more than 30 percent of their work trips are made in the afternoon peak. On the other hand, those persons employed in service industries (such as transportation, personal service, amusement, retail, and professional), which must tailor working hours to their clientele, make a smaller percentage of their work trips in the peak, approximately 20 to 25 percent.

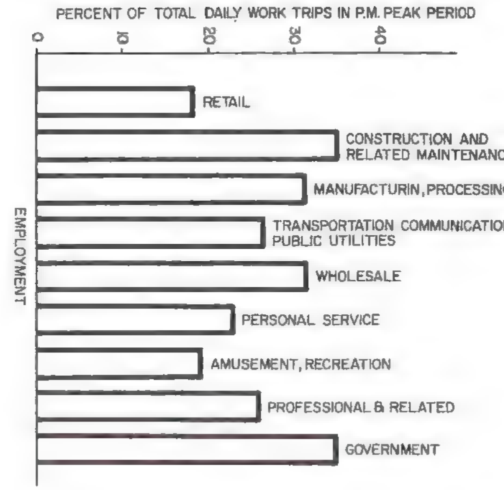


Figure 2. Daily work trips in afternoon peak period by employment of trip-makers.

Although the relationship between work-trip peaking and industry of each zone adequately reproduced attraction-zone peaking, the estimates by production zone were biased, apparently due to the spatial differences in the income of residents. An additional variable was then introduced—the income in the trip-production zone. With this formulation, the model is a trip-interchange relationship treating each i-j trip pair separately.

With this model, there appeared to be 2 distinct types of industries—those in which peaking decreases as income increases and those that illustrate the opposite trend. Office, retail, government, and intensive industries show the former relationships, whereas service, professional, and extensive industries show the latter. This is due to the type of work related to compensation in the respective industries. Higher paid office workers are less likely to have fixed hours (or more freedom to set their own schedules), whereas the reverse is true for those in service industries.

BASE-YEAR CALIBRATION

In some instances, specific model adjustments were required: for the CBD where the trend to uniform working hours leads to somewhat different relationships; for the 2 major employment sites in the region, the Sparrow's Point steel plant with a large third shift and the headquarters of the Social Security Administration that is a specialized office operation; and for other locations, such as hospitals and Friendship Airport, that also have around-the-clock operations.

The final developed relationships for the peaking of work travel are shown in Figure 3. Tests of the model were conducted by using independently derived employment composition data from the Regional Planning Council. On the whole, the model replicated base-year peak-period travel within 1 percent, and less than 3 percent of predicted trip interchanges varied from the origin and destination survey data by more than a single sampled trip.

Figure 4 shows a further independent check—the trip-length distribution of observed and estimated peak-period work trips. This comparison is significant because highway travel time is not a variable in the model.

FORM OF THE MODEL

The model, as developed from the origin and destination data, was of the form

$$P_{it} = f(MF_i, K_t)$$

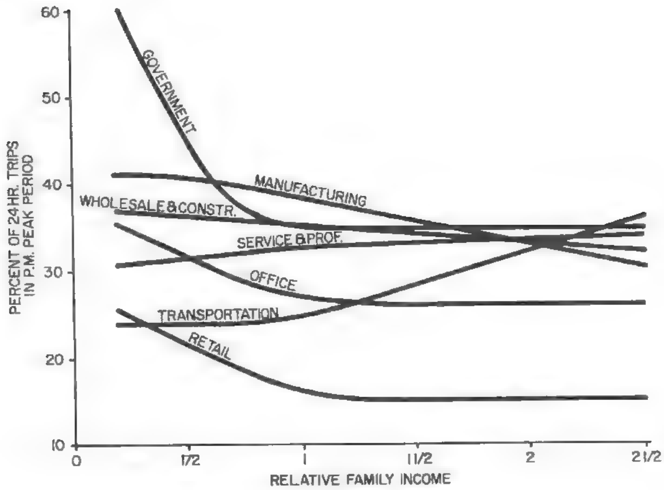


Figure 3. Travel peaking characteristics of non-CBD work trips.

where

P_{it} = percentage of total daily work trips that are produced in zone i by trip-makers employed in industry t, and that occur in the peak period;
MF_i = median family income in zone i; and
K_i = industry in which the trip-maker is employed.

Even though this type of relationship dealing with individual trip-makers can be used in model development, the application of the model to future conditions requires the use of zonal aggregate parameters. Adequate model operation is ensured also by using zonal aggregates in the testing with base-year data. These criteria were met by restructuring the model formulation to

$$P_{ij} = g(MF_i, K_j)$$

where

MF_i = median family income in zone i;
P_{ij} = percentage of total daily trips that are produced in zone i and attracted to zone j and that occur in the peak period; and
K_j = factor expressing the industrial composition of employment in zone j.

The exact formulation is given by

$$P_{ij} = \frac{E_{kj}}{{}_kE_{kj}} \times Q_{Iik}$$

where

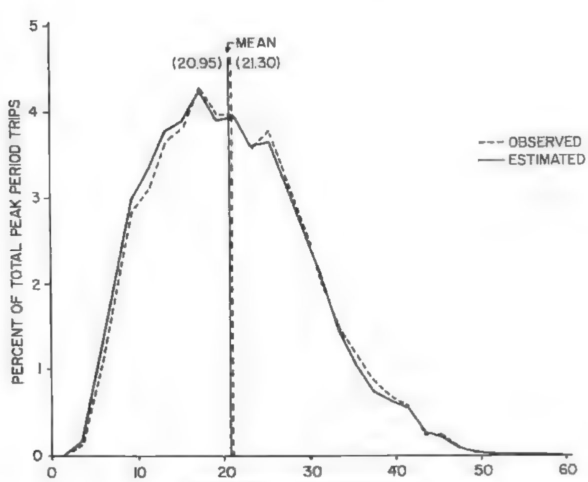


Figure 4. Observed and estimated peak period work trips

P_{ij} = percentage of total daily trips that are produced in zone i and attracted to zone j and that occur in the peak period; and
E_{kj} = number of persons employed in industry k in zone j;
${}_kE_{kj}$ = total employment in zone j;
Q_{Iik} = percentage of trips made in the peak period from zones with median family income I by trip-makers employed in industry k; and
I_i = median family income in zone i.

This equation says, in effect, that the percentage of work trips between zone i and zone j occurring in the peak period can be determined by developing the peaking factor for the median family income in zone i to each industry and then computing an average in which each factor is weighted by the proportion each industry is of the total employment in zone j.

The previously given formulation can best be explained with an example. Assume that there are 2 zones—zone i, which is the production zone (home), and zone j, which is the attraction zone (work)—and that there are 1,000 total daily work trips produced in zone i and attracted in zone j. Further assume the following zonal characteristics:

I_i = median family income in zone i = \$8,000;
E_{1j} = retail employment in zone j = 250;
E_{2j} = government employment in zone j = 200;
E_{3j} through E_{6j} = all other employment in zone j = 0;
${}_kE_{kj}$ = 250 + 200 = 450;
Q_{Ii1} = $Q_{8000,1}$ = percentage of total work trips to retail employment made in the peak period by persons living in zones with median income of \$3,000 = 0.17; and
Q_{Ii2} = $Q_{8000,2}$ = percentage of total work trips to government employment made in the peak period by persons living in zones with median in of \$8,000 = 0.37.

Therefore,

$$P_{ij} = \frac{E_{kj}}{{}_kE_{kj}} \times Q_{Iik}$$

$$= \frac{250}{450} \times 0.17 + \frac{200}{450} \times 0.37$$

$$= 0.095 + 0.164$$

$$= 0.259$$

The percentage of total daytime work trips occurring in the peak period is 25.9 percent and the number of trips is 1,000 x 0.259, or 259 trips. Thus, calibration of the model consists of determining the value of Q_{Iik} for each combination of income and employment type.

NONWORK TRAVEL

The use of industrial composition as an explanatory variable is obviously not valid for nonwork trip purposes. The relationship with income, however, appears to hold for the remainder of the home-based travel. Adult school trips are treated as work trips to a common industry showing a rising percentage of peak travel with rising incomes. For other home-based travel, income seems to be a measure of relative freedom in choosing travel time and of mobility. Those with low and high incomes make other trips in peak hours, perhaps because they are not as likely to be job-holders, whereas the middle-income populations make their shopping and recreational trips outside the peak periods. The relationships for school and other trips are shown in Figures 5 and 6.

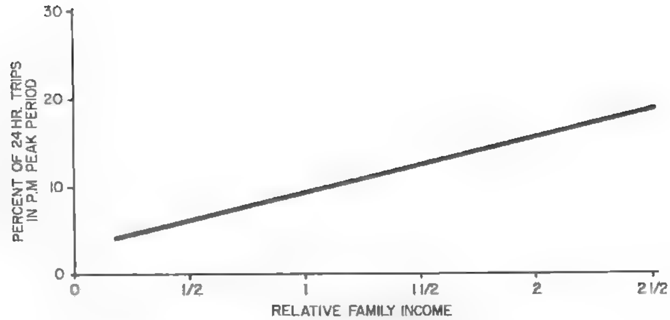


Figure 5. Travel peaking characteristics of school trips.

For non-home-based travel, the peaking percentage showed little variation about a mean value of 16 percent, and further analysis was not deemed necessary.

PEAKING ANALYSIS

A comparison of travel patterns occurring in the peak periods with total daily travel patterns indicates that the orientation of travel in these 2 time periods is significantly different. This difference is due to the wide variation in the type of interaction between various activities in the region. In many cases, the travel from one area to another is made primarily in the peak period; whereas, in other cases, the travel is spread more uniformly over the entire day resulting in less travel, proportionally, in the peak period. This phenomenon depends on the purpose of the trip and social and economic composition of the trip-makers. The highly peaked movements in the afternoon period are from areas that have high employment uses to areas that are more residential in character. This is, of course, quite logical because work-to-home trips are much higher peaked than nonwork trips and, thus, make up the largest share of these movements. The reverse movements, i.e., from residential areas to employment areas, are highly peaked in the morning and therefore much below the average in the afternoon peak period. Movements between 2 areas, both having about the same mix of employment and population, are spread out over the entire day and conform more closely to total daily travel patterns.

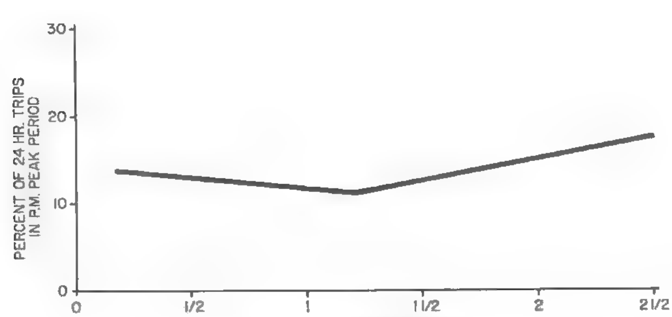


Figure 6. Travel peaking characteristics of other trips.

Method of Analysis

An extensive investigation of all future travel movements between subareas was undertaken to gain a better understanding of the peaking phenomenon. For this discussion, selected subareas, or communities, have been chosen to represent the full range of conditions. The analysis was made by comparing the peaking characteristics of individual communities, as shown in Figure 7, with the average peaking for the entire region.

Peaking From Employment Areas

Figures 8, 9, and 10 show peaking of travel made from the regional core (area 0), the Pulaski-Broening Highway industrial area (area 8), and the Friendship Airport area (area 25). Almost all movements made from these areas are far above the average peaking of 20 percent. A majority of the movements are over twice as peaked as the average. Across-the-board factoring of total daily travel movements would understate these movements by more than a factor of two. Only trips made from the Friendship Airport area to other more predominant employment areas could be described as average or below average. All other movements are at least 50 percent above average. Although not illustrated, similar peaking was found for travel generated from other employment areas in the city and counties, e.g., those with an employment-to-population ratio of 0.5 and more. Table 1 gives the average peaking of all travel made from

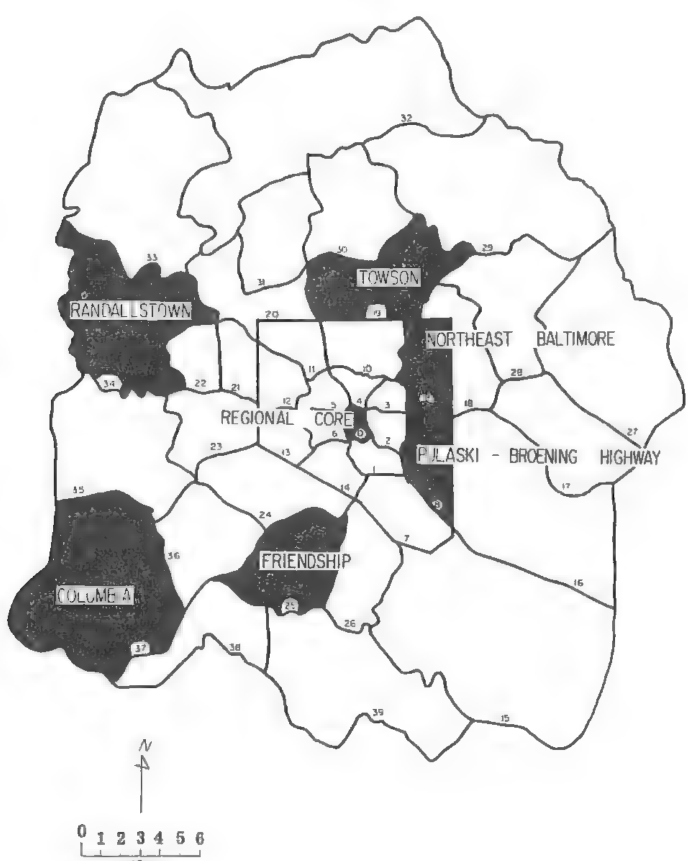


Figure 7. Communities selected to illustrate peaking characteristics.

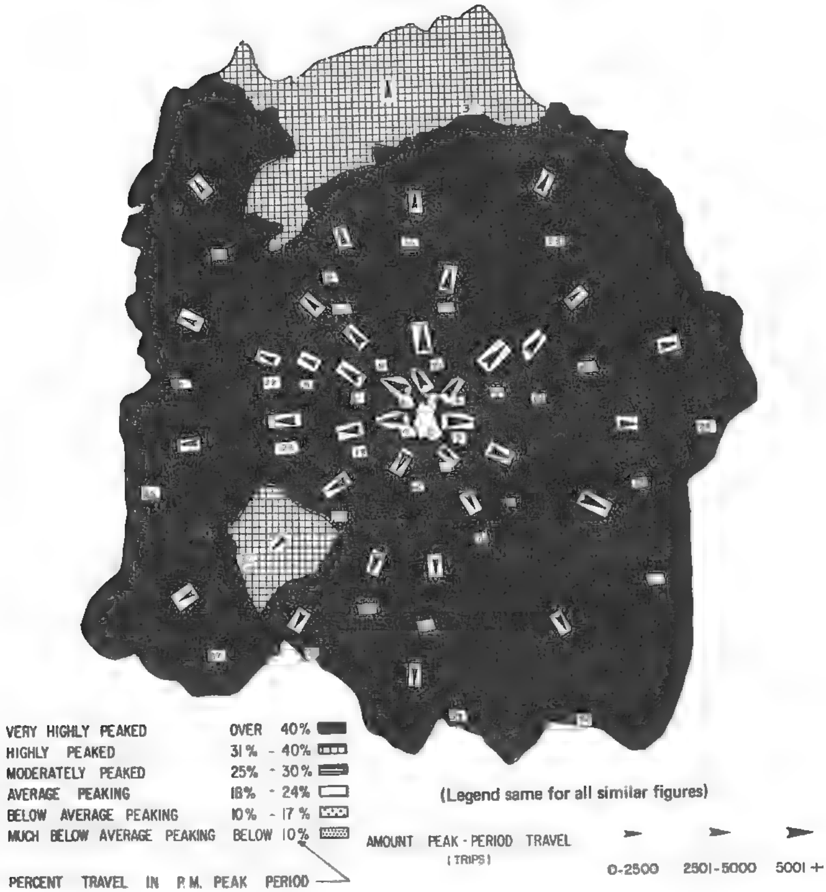


Figure 8. Peaking of 1990 travel leaving the regional core, a dominant employment community.

communities that are primarily employment areas. As can be seen, a great amount of travel would be understated by applying a uniform factor to total daily travel.

Peaking to Employment Areas

Peaking patterns of trips going to the same 3 selected employment areas are shown in Figures 11, 12, and 13. As would be expected, the reverse movements to these high employment areas are far below average in the afternoon peak period. Total daily movements factored by a uniform factor would overstate these movements by a significant amount. However, some trips made from certain city industrial areas to the Friendship Airport area are highly peaked in the afternoon because the number of workers heading to their homes in the Friendship Airport area exceeds the number of workers heading to homes in these city areas. Therefore, there is a preponderance of flow toward Friendship Airport in the afternoon.

Peaking to Residential Areas

Two communities, northeast Baltimore (area 9) and Randallstown (area 34), were selected to illustrate peaking of travel destined to residential areas. Many other communities in both the city and counties are also primarily residential and have similar

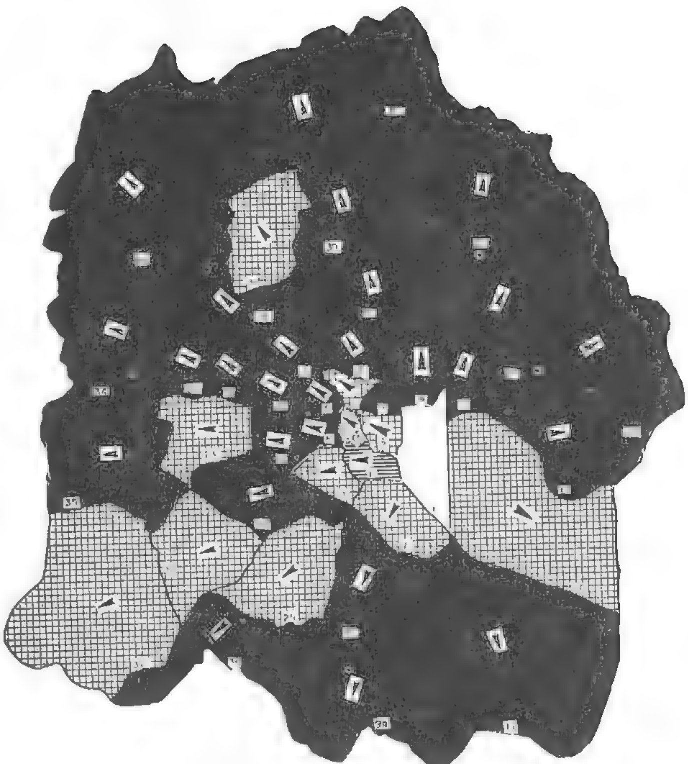


Figure 9. Peaking of 1990 travel leaving the Pulaski-Broening highway area, a dominant employment community.

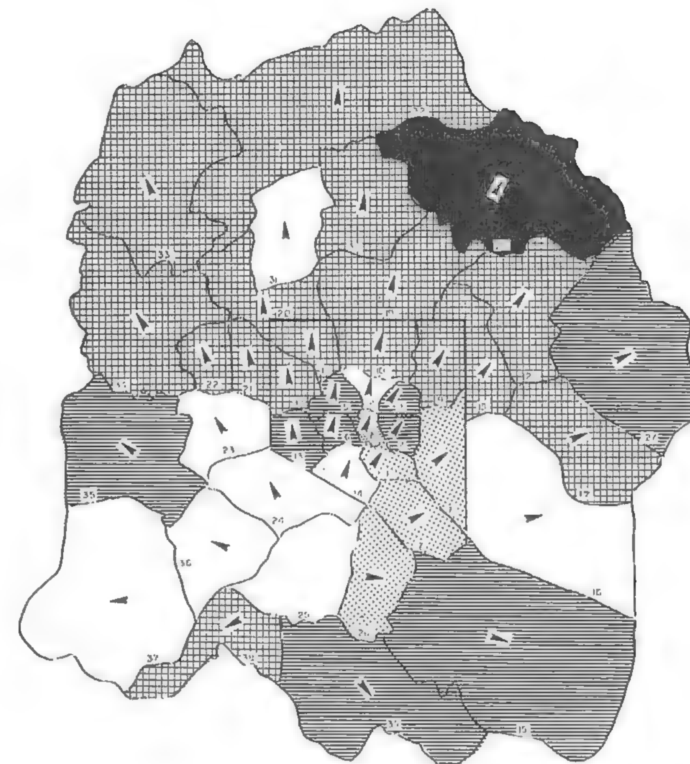


Figure 10. Peaking of 1990 travel leaving the Friendship Airport area, an area of high employment.

TABLE 1

AVERAGE PEAKING OF ALL TRAVEL MADE FROM COMMUNITIES THAT ARE PRIMARILY EMPLOYMENT AREAS

Area, Designation	Employment to Population Ratio	1990 Daily Origins	Peak-Period Origins Based on Uniform Factor[a]	Simulated Peak-Period Origins	Difference
Regional core, 0	4.5	258,500	51,700	112,900	61,200
Pulaski-Broening Highway, 8	1.6	107,000	21,400	39,600	18,200
South Baltimore, 1	1.0	28,600	5,700	10,200	4,500
Curtis Bay, 7	1.0	29,300	5,900	10,100	4,200
Hampden-Waverly, 4	0.9	81,200	16,200	21,600	5,400
Friendship Airport, 25	0.8	85,800	17,000	18,800	1,800
Social Security Administration, 23	0.7	142,500	28,600	41,600	13,000
Dundalk-Sparrows Point, 16	0.5	164,900	33,000	34,800	1,800
Eastern Howard, 36	0.5	53,600	10,700	13,200	2,500
East Baltimore, 2	0.5	102,700	20,500	26,800	6,300
Cherry Hill-Lakeland, 14	0.5	69,400	13,900	16,200	2,300
Total		1,123,500	224,600	345,800	121,200

[a]Uniform factor of 20 percent, i.e., average peaking factor for all travel in metropolitan area.

Figure 11. Peaking of 1990 travel going to the regional core, a dominant employment community.

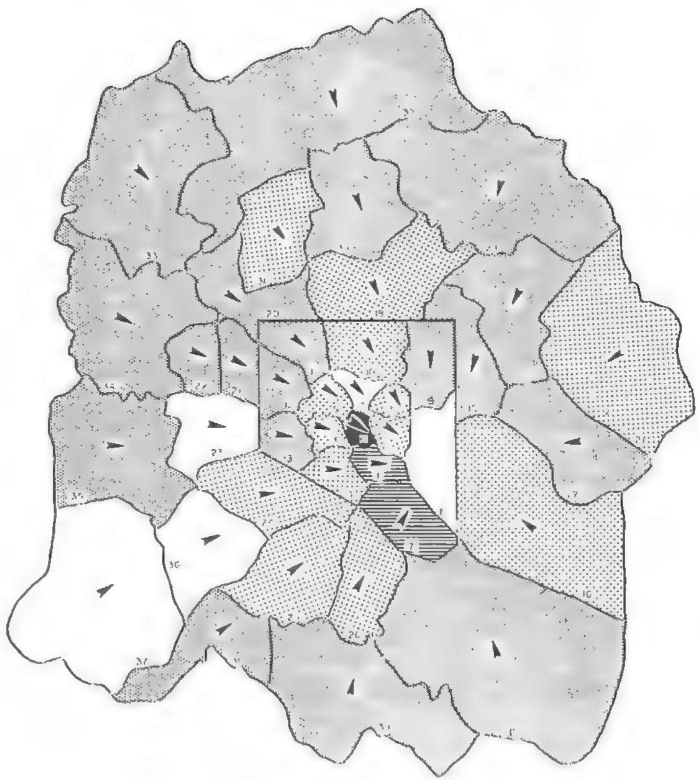


Figure 12. Peaking of 1990 travel going to the Pulaski-Broening Highway area, a dominant employment community.

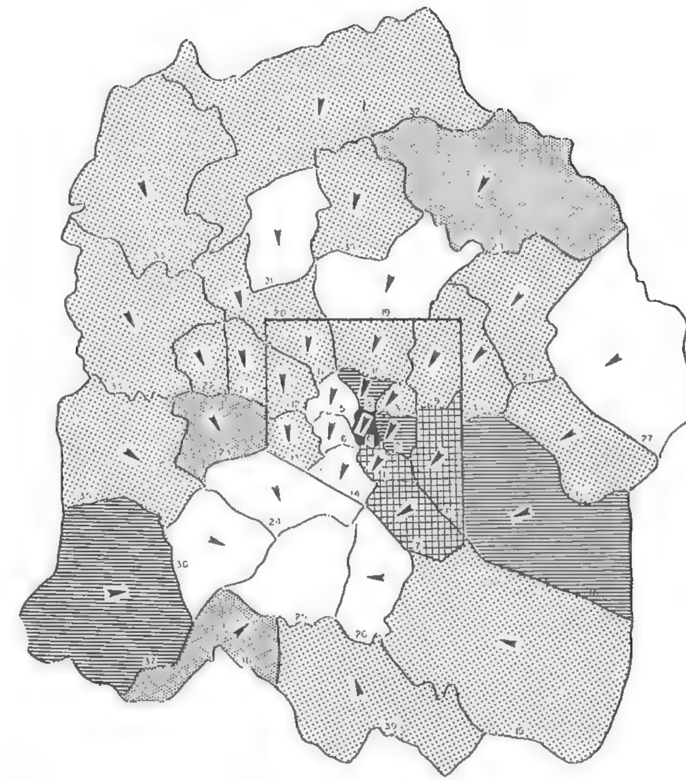


Figure 13. Peaking of 1990 travel going to the Friendship Airport area, a high employment community.

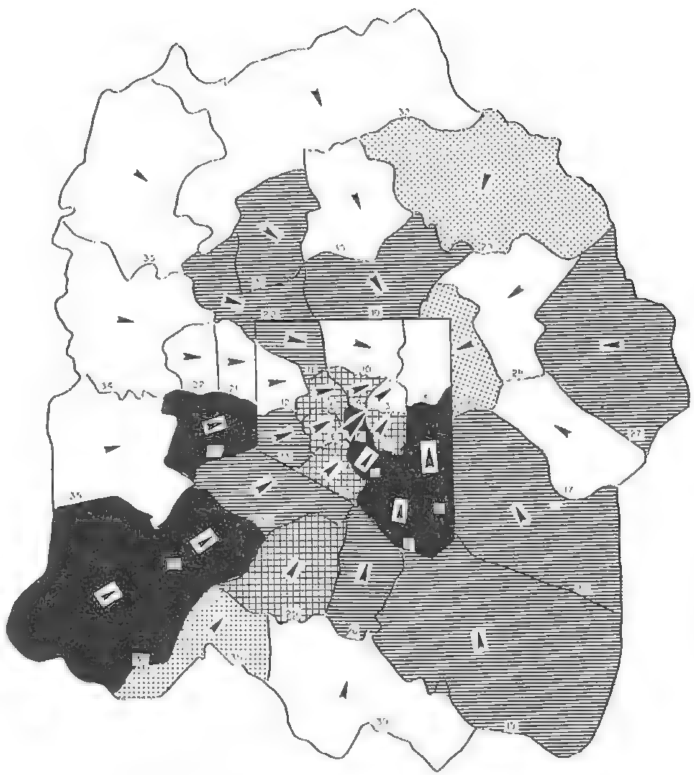


Figure 14. Peaking of 1990 travel going to northeast Baltimore area, a dominant residential community.

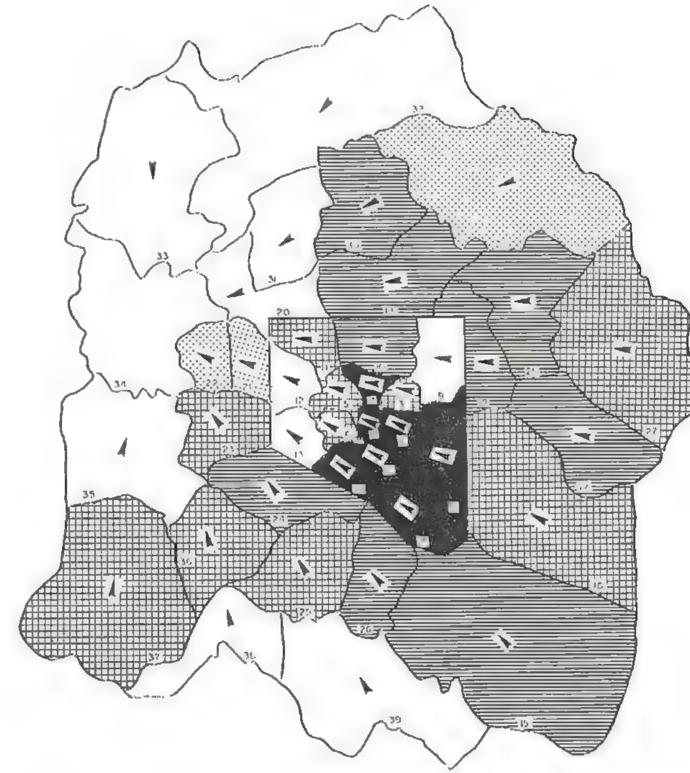


Figure 15. Peaking of 1990 travel going to the Randallstown area, a dominant residential community.

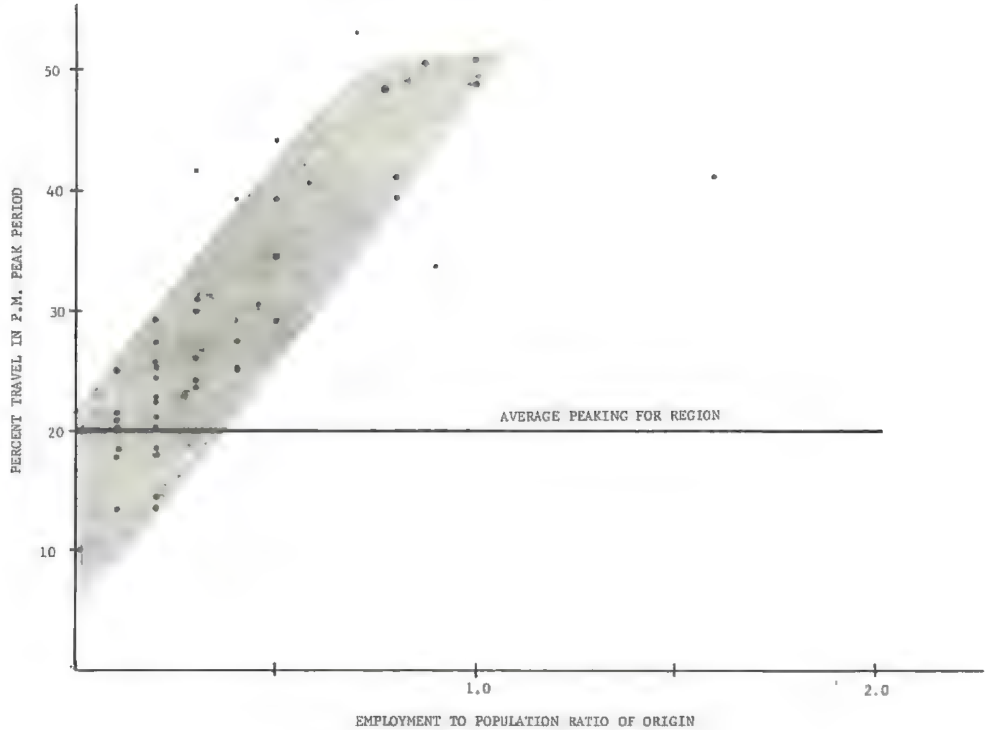


Figure 16. Peaking of travel going to northeast Baltimore as related to the employment and population mix of the origin.

peaking patterns. An examination of Figures 14 and 15 clearly shows that the degree of peaking to these areas depends, to a large degree, on where the travel originates. As shown earlier, travel from predominant employment areas is highly peaked, whereas travel from other residential areas more nearly conforms to the average. Figure 16 shows the relationship between peaking of travel destined to northeast Baltimore and the employment and population mix of the origins of the travel. Logically, peaking of travel is highest from high employment areas and is consistently less from areas more residential in use.

Peaking From Residential Areas

Figures 17 and 18 show peaking of travel destined to the same 2 residential areas. These patterns form a "reverse mirror image" of peaking patterns of travel originated in these areas. Where peaking is high for trips from employment areas, the reverse movement is low.

Peaking to and From Mixed Areas

Both Towson (area 19) and Columbia (area 37) conform to the average mixture of employment and population of the region. Peaking of travel to and from these communities is shown in Figures 19, 20, 21, and 22. The degree of peaking varies depending on the specific movement. Figure 23 shows how peaking of movements going to Towson varies as a function of the employment-to-population ratio of the origin of the trip. This relationship appears to be quite logical.

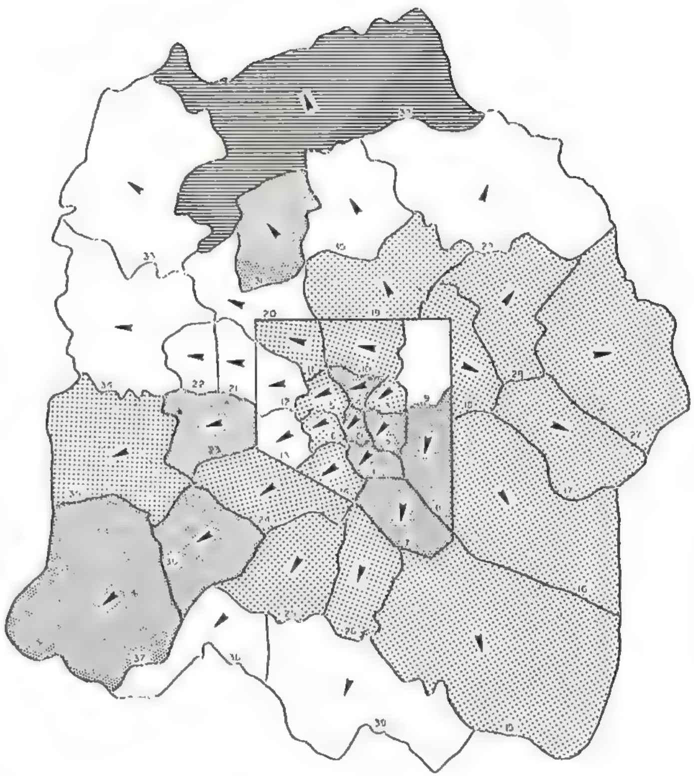


Figure 17. Peaking of 1990 travel leaving northeast Baltimore, a dominant residential community.

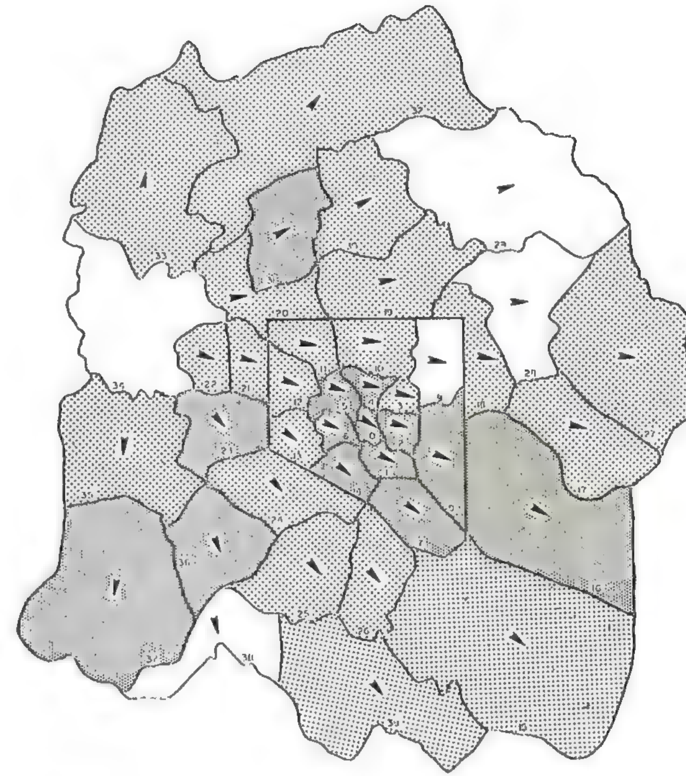


Figure 18. Peaking of 1990 travel leaving the Randallstown area, a dominant residential community.

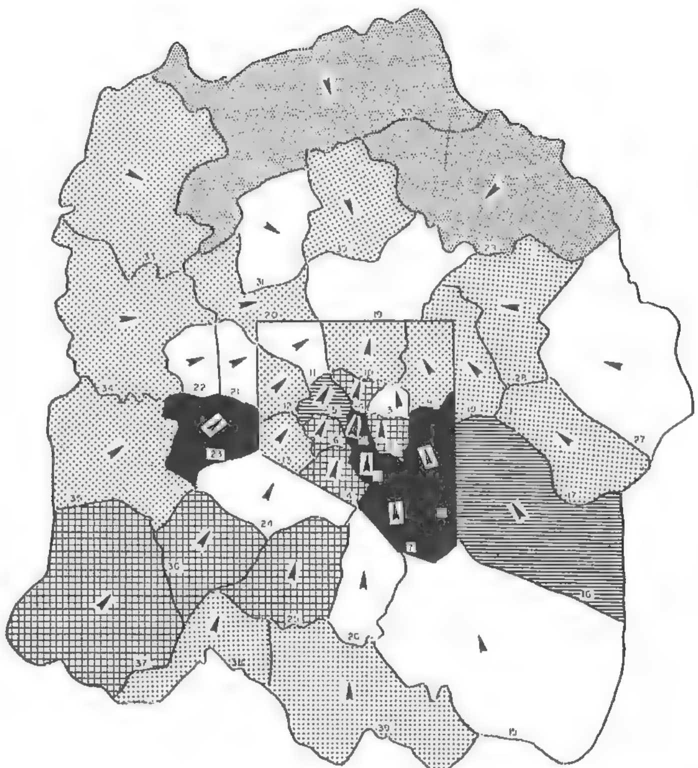


Figure 19. Peaking of 1990 travel going to the Towson area, a mixed community.

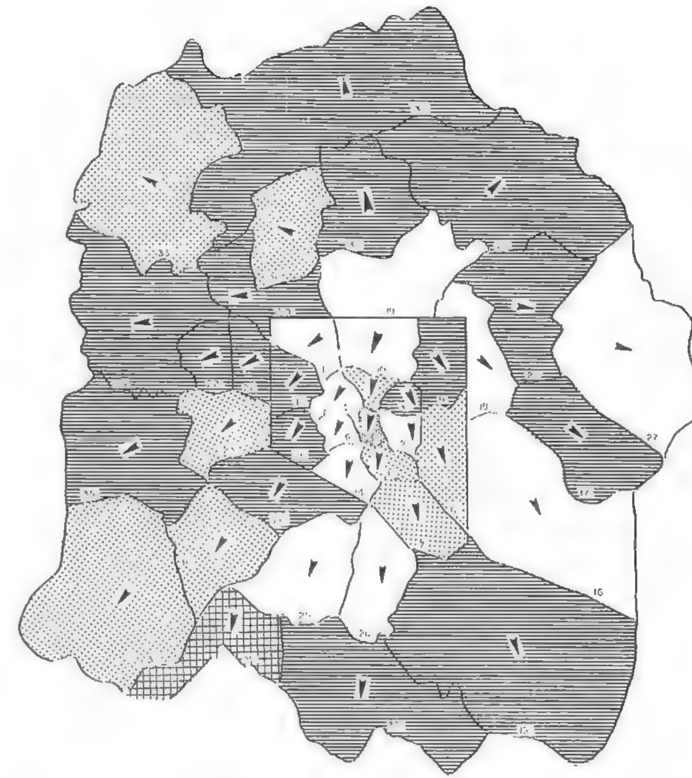


Figure 20. Peaking of 1990 travel leaving the Towson area, a mixed community.

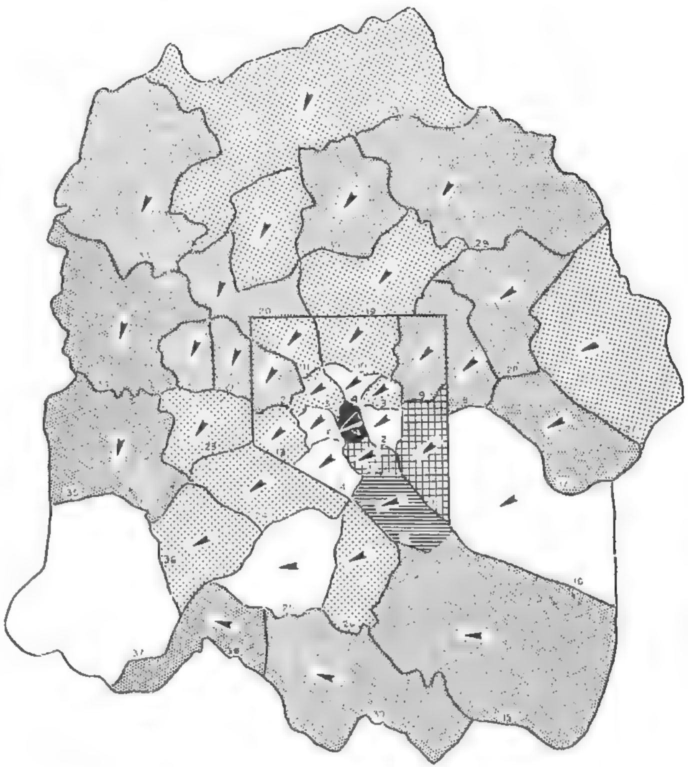


Figure 21. Peaking of 1990 travel going to the Columbia area, a mixed community.

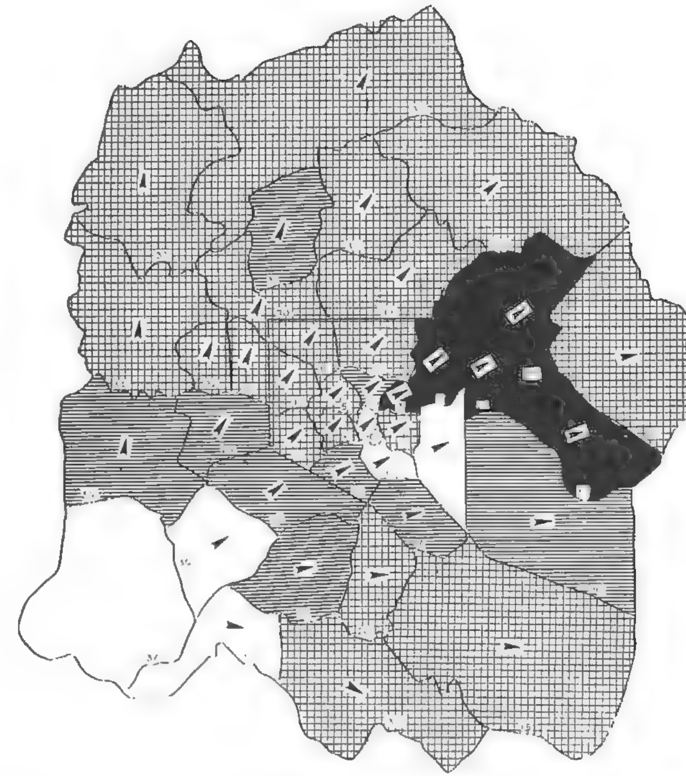


Figure 22. Peaking of 1990 travel leaving the Columbia area, a mixed community.

TABLE 2

AMOUNT OF INTERCOMMUNITY PEAK TRAVEL THAT FALLS INTO EACH PEAKING CATEGORY

Peak Category	Percent of Travel	Peak-Period Trips	
		Number	Percent
Very highly peaked	Over 40	156,000	21
Highly peaked	31 to 40	91,000	12
Moderately peaked	25 to 30	92,000	12
Average peaking	18 to 24	242,000	33
Below average peaking	10 to 17	136,000	19
Much below average peaking	below 10	22,000	3
		739,000	100

Note: Data include all internal travel made in the metropolitan area in 1990.

Overall Peaking Variation

Clearly, peaking of specific travel movements varies greatly depending on the composition of the movement. Table 2 gives the amount of intercommunity peak travel that falls into each peaking category. Only about one-third of the travel conforms to the average regional peaking percentage. Nearly 50 percent of the travel can be described as being significantly peaked (more than 25 percent above the average). More

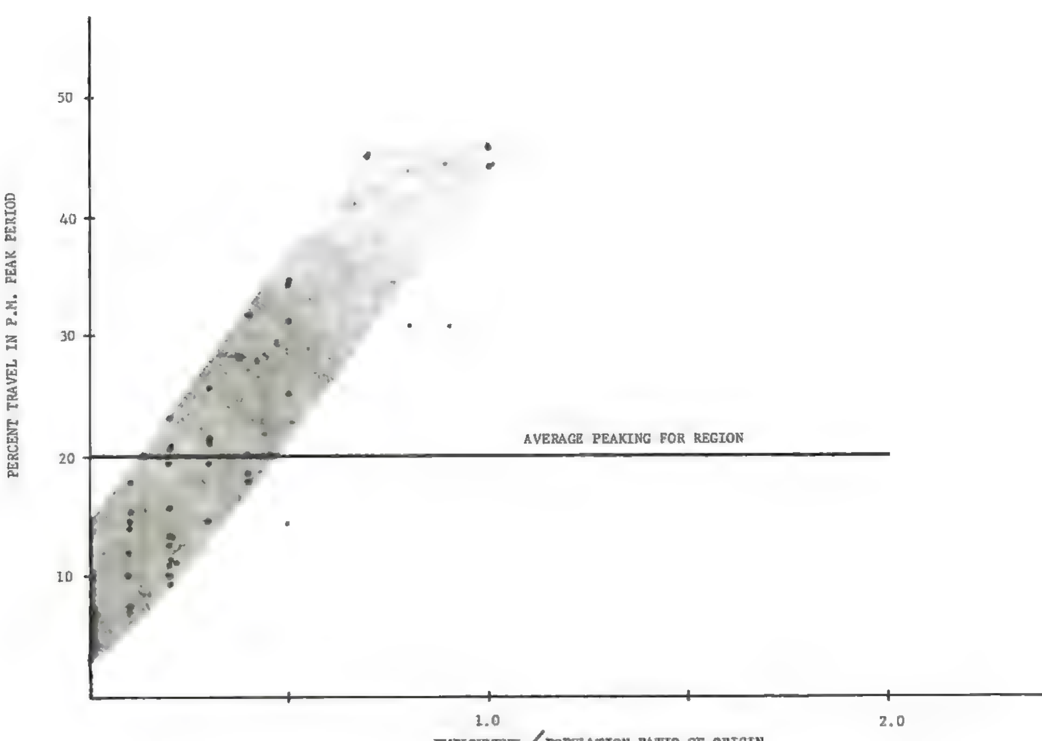


Figure 23. Peaking of travel going to Towson as related to the employment and population mix of the origin.

TABLE 3

AMOUNT OF ERROR RESULTING FROM USE OF 1962 PEAKING FACTOR TO ESTIMATE 1990 PEAKING

Community Number	1962 Afternoon Peak-Period Origins (percent)	1990 Daily Origins	Peak-Period Origins Based on 1962 Factor for Each Community	Simulated Peak-Period Origins	Difference[a]
0	47	258,500	121,200	112,900	+8,300
1	30	28,600	8,600	10,200	-1,600
2	23	102,700	23,600	26,800	-3,200
3	15	53,300	8,000	8,800	-800
4	32	81,200	26,000	21,600	+4,400
5	14	94,300	13,200	18,300	-5,100
6	17	78,600	13,300	18,700	-5,400
7	25	29,300	7,300	10,100	-2,800
8	30	107,000	32,000	39,600	-7,600
9	10	143,300	14,300	17,600	-3,300
10	15	138,500	20,800	21,700	-900
11	15	110,400	16,600	19,600	-3,000
12	11	67,800	7,500	7,700	-200
13	13	78,500	10,100	11,100	-1,000
14	31	69,400	21,500	16,200	+5,300
15	10	103,700	10,400	14,500	-4,100
16	14	164,900	23,100	34,800	-11,700
17	10	79,300	7,900	10,200	-2,300
18	11	103,300	11,400	14,000	-2,600
19	19	211,400	40,200	42,100	-1,900
20	19	69,300	13,200	11,000	+2,200
21	9	43,500	3,900	5,100	-1,200
22	11	44,900	4,900	5,200	-300
23	19	142,400	27,000	41,600	-14,600
24	15	122,500	18,400	19,400	-1,000
25	27	85,800	23,200	18,800	+4,400
26	13	125,300	16,300	22,100	-5,800
27	36	68,900	24,800	12,100	+12,700
28	11	106,400	11,700	14,800	-3,100
29	12	39,400	4,700	4,100	+600
30	16	74,400	11,900	10,600	+1,300
31	18	32,800	5,900	6,700	-800
32	12	23,900	2,900	2,700	+200
33	19	97,300	18,300	14,500	+3,800
34	10	102,200	10,200	11,600	-1,400
35	14	66,300	9,300	8,900	+400
36	19	53,600	10,200	13,200	-3,000
37	16	116,100	18,600	27,800	-9,200
38	16	29,100	4,600	2,700	+1,900
39	9	83,300	7,500	9,800	-2,300
Total	19	3,631,400	684,500	739,200	-54,700

[a]A plus sign indicates that trip-making would be overestimated with 1962 peaking factor, whereas a negative sign indicates that trip-making would be underestimated.

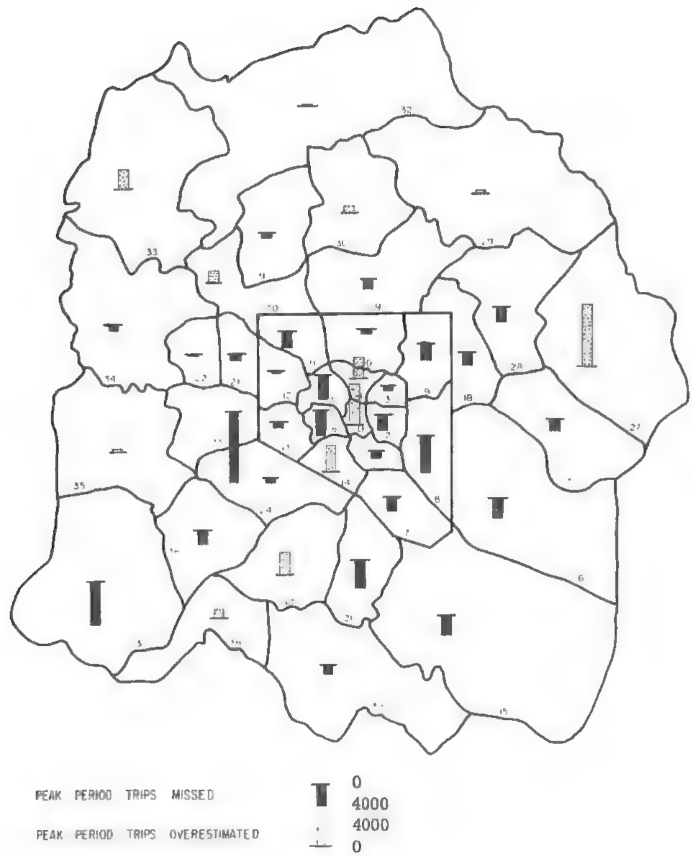


Figure 24. Amount of error resulting by using 1962 peaking factor on 1990 travel origins.

than 20 percent of the travel, mostly radiating out of employment areas, is very highly peaked.

Peaking Changes Through Time

Travel simulation techniques make it possible to account for changes in travel patterns through time. An analysis of projected peaking changes between 1962 and 1990 indicates that estimates of peak-period travel based on base-year observations cannot provide reasonably accurate estimates of future peak-period conditions. Peaking characteristics were found to change through time as a result of the uneven growth in employment and population. Major errors would result by applying peaking percentages existing in the base year to future total daily travel patterns.

Table 3 gives a comparison, by community, of the simulated 1990 peak-period travel origins with 1990 24-hour origins factored by the 1962 peaking factor for each community. Overall, nearly 55,000 trips in a 2-hour period would be missed. The resulting errors of travel made from most communities could be very serious. These errors, as shown in Figure 24, would vary depending on the type of growth expected. For instance, in 1962 about 36 percent of the travel made from the Middle River community (area 27) was made in the afternoon peak period. This high peaking was primarily due to the large employment in the area in proportion to the population residing in the area. By 1990 this is estimated to drastically change. Employment in 1990 is projected to be

less than in 1962 because of the major employment drop at Martin-Marietta, whereas the population will grow significantly. Therefore, the peaking factor will decrease to 18 percent. Application of the base-year peaking factor to future total day travel would overestimate future peak-period travel by nearly 13,000 trips in the 2-hour period. Another example is the area of the Social Security Administration (area 23). In this case peaking will become more pronounced through time as a result of the rapid employment growth expected at the Social Security Administration. The peaking factor in 1962 was about average at 19 percent. In 1990, it is expected to be nearly 30 percent. Applying the 1962 factor to the 1990 forecast of total daily traffic would underestimate the 1990 peak-period demand by nearly 15,000 trips.

Based on these analyses, it is clear that peaking factors will change through time and that the use of peak-period simulation model will make it possible to avoid serious errors in estimating future level of service.

CONCLUSIONS

The greatest component of travel in the peak period is the trip from the place of employment to home. As discussed previously, the peaking characteristic of this type of trip is dependent on the income of the trip-maker and the industry in which he is employed. Further, there appear to be 2 distinct types of industries showing differing relationships with income, based on the relative position of white-collar workers in these industries.

As would be expected, those industries that are fairly well self-contained and have standard working hours, such as government, manufacturing, and construction, show a greater percentage of travel in the peak period than do industries that must adjust their working hours to customer demand, such as service, retail, and recreation.

The amount of travel in future peak periods, then, will be mainly a function of the industrial composition of the region. Increased employment in the governmental sector could cause more severe peaking problems than found today, particularly in areas with concentrations of government workers. On the other hand, if there is a shift of workers into service and retail categories away from manufacturing and construction, then there may well be a substantial drop in peak-period travel demand resulting in more efficient utilization of transportation facilities.

The school relationship as developed is a linear function increasing with increasing income. Because school bus travel is not included, the model relates mainly to higher educational institutions. The model suggests that the higher the income is, the greater the percentage of home-based school trips will be in the peak period. Although upper income groups probably do make more trips to educational facilities, low-income people also attend these schools but quite frequently hold a job while attending school. Their trips from jobs to school (or vice versa) are categorized as non-home-based trips and thus do not appear in the school peak-period model. However, as incomes and greater educational participation increase, more school travel seems likely to occur in the peak period in future years.

The relationships for other trips state that the degree of peaking for these trips declines slightly as income increases up to a specific income, $6,500 for the Baltimore Region, and then increases as income increases further. This is logical because the lower a person's income is the more restricted he is as to when he can make a trip. Moreover, he usually will not make a trip in the other category unless it is absolutely necessary. As income increases above the $6,500 level, people start making trips that are associated with more affluence. These trips include those by women coming home from club meetings or picking up children from some after-school function. This indicates that as incomes rise and more leisure time becomes available, more other trips will occur in the afternoon peak period.

Travel not originating from the home is difficult to quantify, and the peak period apparently is a common characteristic of this type of trip throughout the region. This lack of variation, illustrated by the use of a mean value to reproduce the peak-period travel pattern, indicates that little change can be expected in this group in future peak periods.

The comparison of model projections against extrapolation of present trends clearly indicates that a model that is sensitive to changes in the distribution and composition

of urban activities is needed to properly project future peak-travel demand as part of evaluation and design of transportation systems.

REFERENCE

1. 1975 Transportation Plans. Penn-Jersey Transportation Study, Vol. 3, May 1965.

SYNTHESIS OF VEHICLE TRIP PATTERNS IN SMALL URBAN AREAS

Hatim M. Hajj, Harland Bartholomew and Associates

This paper presents the development of a simplified technique for trip generation and distribution for the small urban area of Madisonville, Kentucky. This model, developed on the basis of experience in small urban areas, is an economical alternative to an internal origin and destination study. The input to the study consists of an external origin and destination survey, existing daily traffic volumes on the major street system, and limited socioeconomic data. Internal vehicle trips and corresponding travel time factors were developed based on experience in small urban areas. Internal and external trips were related to population and employment through a multiple linear regression procedure. The conventional gravity model was used to distribute trips. Analytical and statistical tests were used in the calibration of the model. The ultimate test was for the model to reproduce as closely as possible the ground counts on the major street system. The conclusions of the study were that the developed model adequately synthesized internal trip patterns in the Madisonville urban area. The model uses 3 socioeconomic factors that can be easily forecast—population, total employment, and industrial employment. Such a procedure is recommended for use in small urban areas. However, the model needs to be developed and calibrated for each small urban area.

•IN RECENT YEARS, a major emphasis in transportation planning in urban areas has been development of procedures and methodology for solving planning problems in large urban areas. The conventional comprehensive origin and destination surveys with the use of home interviews to collect existing travel data often cannot be economically justified for small urban areas (population less than 50,000). Synthesis of existing travel patterns can provide a feasible alternate to the conventional comprehensive origin and destination surveys. In addition, typical transportation study models require data for many independent variables, yet comprehensive inventories of all of these items often cannot be warranted for studies in small urban areas.

One of the prerequisites for detailed forecasts of travel patterns is a conceptually reasonable, operationally realistic, and financially feasible model for trip generation and trip distribution. This paper presents such a model that was developed for synthesizing trip patterns in the Madisonville, Kentucky, urban area. The synthesis procedures provide a means for determining existing travel patterns and for forecasting future trip patterns. Moreover, the model was 3 socioeconomic factors that can be easily projected—population, total employment, and industrial employment.

The Madisonville urban area is situated in Hopkins County, Kentucky, 110 miles southwest of Louisville and 60 miles northeast of Paducah. In 1968, the study area population was 18,224, and the total employment in the study area was 6,036.

An external origin and destination survey was conducted for the Madisonville study area in 1968. At the same time, surveys of a planning-inventory nature were performed to provide data for use in all phases of the comprehensive plan. The study area was divided into 50 analysis zones. Eight external stations were established. The results

Sponsored by Committee on Transportation Forecasting and presented at the 50th Annual Meeting.

of these inventory surveys provide the base data regarding characteristics of the area and existing travel patterns within the area on a typical day. These data are needed for model preparation.

METHOD OF APPROACH

The approach used in the Madisonville Urban Area Transportation Study (MUATS) has been to develop an overall traffic model that will adequately predict the existing travel pattern within the study area on a typical weekday. Development of this model is composed of 2 distinct phases—trip generation and trip distribution. These phases incorporate the use of mathematical expressions or equations that are based both on observations of travel within the study area and on various travel-related characteristics of the study area. A general outline of the model development procedures utilized in the Madisonville study is given in Table 1 and is described in the following paragraphs.

In the trip-generation phase of the model development process, generation is defined as the sum of trips beginning and ending in an area. In the development of trip attractions and trip productions, the traditional trip-purpose concept of home-based trips and non-home-based trips is followed. For home-based trips, a trip production is defined as the home end of the trip, which can be either a trip origin or a trip destination. A trip attraction is defined as the nonhome end of the home-based trip, and it also can be either a trip origin or a trip destination. In the non-home-based category of trips, the origin is always the production end, and the destination is always the attraction end.

Because an internal origin and destination survey was not available for the Madisonville study area, the internal trip-attraction and trip-production equations were developed by using the results of comprehensive origin and destination studies in other small urban areas. In the development of the mathematical model for internal and external trips, the data from the external origin and destination survey conducted for the Madisonville urban area were summarized for each traffic zone. These observations were then mathematically related to a number of known conditions within each zone, such as population, dwelling units, and employment. From the series of mathematical expressions analyzed, the best estimating equation was chosen for use in calculating both the existing and the future internal and external trip attractions. The selection of the appropriate internal and external trip-attraction equation was based on a consideration of the possibility that travel characteristics different from those existing at the time of the external origin and destination survey might occur.

The trip-distribution analysis involves a determination of the number of trips that will be produced in 1 zone and attracted to another zone. The conventional gravity model was used for the distribution of trips among zones. This trip-distribution process requires trip-production and trip-attraction data for each zone in the study area and a measure of the spatial separation between each pair of zones in the study area. Thus, the need for compatibility between the trip-generation analysis and the trip-distribution analysis becomes apparent.

The following objectives are important in the development of a mathematical traffic model:

1. Establishment of associations between model elements (trip generation and trip distribution) and the physical development of the study area;
2. Maintenance of control of the distribution rates for trips with different trip lengths;
3. Development of traffic-model procedures compatible with a total model concept (direct progression from an updated planning data file to estimated traffic on the streets); and
4. Maximization of the adaptability of the results to continuing planning activities.

A number of validity checks (statistical and analytical) are made to ensure acceptability of a traffic model. The basic philosophy involves an acceptance of the model if it reproduces the ground counts to an adequate degree. In general, the primary purpose of the validity checks is to ensure the best possible match between the total traffic volume on each segment of the major street system as derived from the models and as measured in the field. Discrepancies between these 2 values are resolved until the

TABLE 1

OUTLINE OF SIMULATION PROCEDURE (MUATS)

I. Trip generation

A. Internal trips

1. Determine internal vehicle trip rate per capita based on experience in urban areas of less than 50,000 population.
2. Determine percentage distribution among trip purposes (home-based work, home-based nonwork, and non-home-based) based on experience in relatively small urban areas.
3. Select pertinent trip generation equations by reviewing mathematical models for small urban areas.
4. Determine existing internal zonal productions and attractions by purpose by using selected trip-generation equations cross-checked with reasonable trip rate per capita and reasonable percentage distribution among trip purposes.
5. Balance attractions to productions for each trip purpose. Adjust if necessary.

B. Internal and external trips

1. Develop multiple linear regression model based on trip data from external origin and destination survey.
2. Use population, total employment, dwelling units, commercial employment, industrial employment, and public employment as the only independent variables.

II. Existing major street network

A. Prepare existing major street network according to new Federal Highway Administration procedures.
B. Assign speeds to the various facilities based on functional class and location within the urban area.

III. Trip distribution

A. Internal trips

1. Determine travel time factors for each trip purpose based on experience in relatively small urban areas.
2. Distribute productions and attractions from step 1 by using selected travel time factors and the existing major street network in accordance with gravity model procedures.

a. Use 1-min terminal time in residential zones, 2 min in industrial areas, and 2 to 3 min in commercial areas (3 min in the CBD).
b. Check resulting trip length frequencies for reasonableness.

B. Internal and external—Develop the internal and external travel time factors from the existing trip patterns as obtained from the external origin and destination survey, in accordance with standard gravity model procedures.

IV. Traffic assignment

A. Assign the internal and external trips, memory J, to the existing major street network on an all-or-nothing basis.
B. Assign the total trips, memory J, (internal and external trips, through, and distributed internal trips) to the existing major street network on all-or-nothing basis.
C. Plot assigned volumes and ground counts on a major street network map.

V. Validity checks

A. Compare total assigned volumes with most recent ground counts for the major streets.
B. Compare total vehicle-miles by street class for the synthetic data and the ground counts.
C. Screenline comparisons

1. Establish north-south and east-west (general orientation) screenlines.
2. Establish loop screenlines around the CBD and surrounding area.
3. Compare total volumes at each screenline segment for the synthetic data and ground counts.

D. Trip comparisons and link volumes—Analyze frequency distribution of differences between ground counts and assigned volumes. For each volume group, determine the root mean square error. Apply this test to the final calibration of the models only.

VI. Model adjustments

A. Adjust the models according to results of comparisons (step V) between synthetic data and ground counts.
B. Adjust travel time factors if total vehicle-miles agree reasonably but the distribution by location is off.
C. Adjust the zonal productions and attractions up or down if total number of trips is not being simulated.
D. Adjust both travel time factors and the productions and attractions if both the total vehicle-miles and the distribution by location are off.
E. Adjust speeds so that model is brought as much as possible to link-by-link agreement with ground counts.

differences are reduced to levels that are not significant in terms of the overall volume of travel within the study area.

A review of the mathematical model development and a consideration of the application of the model suggest that the following factors are pertinent:

1. The models are derived from data collected in the external origin and destination survey and from data available in planning information files, and therefore the model results can be no better than the base data.
2. In the final analysis, the forecast of travel is reduced to a lane of traffic. The service volume of a lane of traffic varies from 400 to over 1,800 vehicles per hour, depending on the type of facility and the prevailing physical conditions.
3. The data obtained from the application of mathematical models in long-range planning studies are used to guide the decision-making process of the professional transportation planner rather than to dictate decisions. The model results are guides to future levels of service that will be desired by the public for movement of persons and goods among future concentrations of activities within the area considered. Furthermore, the professional's recommendations are guides for those public officials responsible for implementation of the improvement programs needed to accommodate the travel desires of the visitors to and residents of the Madisonville area.

PRIMARY DECISIONS IN MODEL DEVELOPMENT

A number of initial decisions must be made prior to undertaking detailed analyses in the development of mathematical models, and each of these decisions usually requires some reevaluation during the analysis. The primary decisions dealing with the mathematical model development for the Madisonville Urban Area Transportation Study were as follows:

1. Distribution model—It was decided to develop a 24-hour vehicle trip gravity model because of the small size of the urban area.
2. Trip purpose classifications—In view of the relatively small number of total trips anticipated and the limited planning data available for use, the trip categories selected for the model development were as follows: (a) home-based work, automobile; (b) home-based nonwork, automobile; (c) non-home-based, automobile and truck; and (d) internal and external, automobile and truck.
3. Treatment of external trips—The internal and external trips include those trips made by residents of the study area and those made by nonresidents of the study area. The initial decision on these trips was to combine the local trips (those trips made between the study area and an external area) regardless of whether the trips were made by residents or nonresidents. The traffic zone containing the internal end of these trips was considered to be the attraction, and the external station was considered to be the production. Trips with neither end in the study area (through trips) were to be handled by a factor-expansion method.

TRIP-GENERATION ANALYSIS

Internal Trip Purposes

Because of the lack of internal origin and destination survey data (i.e., home-interview and truck and taxi survey data), all internal trip-generation information had to be synthesized. The synthesis of internal trips was based on the assumption that travel patterns are similar in areas having similar population and socioeconomic characteristics. This synthesis involved a review of mathematical models previously developed for a number of other small urban areas having a population of less than 100,000.

Because of the relatively small number of planning data factors available for use in the trip-generation analysis for the MUAT study, the number of other small urban areas that could be included in this review was limited. The planning data available for the MUATS study consisted of population, dwelling units, total employment, industrial employment, commercial employment, and public employment. Thus, to develop an initial set of generation equations, primary emphasis in this review was placed on delineation

TABLE 2

TOTAL INTERNAL VEHICLE TRIPS PER CAPITA

Study Area	Base Year	Population	Vehicle Trips Per Capita
50,000 Population or Less			
Murray, Kentucky	1967	10,917	2.80
Joplin, Missouri	1960	16,300	2.93
Clarksville, Tennessee	1965	10,993	2.29
Gainesville, Florida	1960	21,351	1.94
Pittsburg, Kansas	1961	22,476	2.38
Staunton, Virginia	1965	32,200	1.69
50,000 to 100,000 Population			
Pine Bluff, Arkansas	1965	67,000	2.00
Fort Smith, Arkansas	1965	80,000	1.83
Kingsport, Tennessee	1966	67,000	2.37
Tyler, Texas	1964	64,512	2.38
Sioux Falls, South Dakota	1963	83,684	2.45
Rapid City, South Dakota	1963	73,459	1.87
Great Falls, Montana	1961	74,735	1.73
St. Joseph, Missouri	1962	84,169	1.99
Appleton, Wisconsin	1964	54,100	1.92
Laredo, Texas	1964	64,311	1.57
San Angelo, Texas	1964	63,438	2.16
Texarkana, Texas	1964	64,278	1.64

of other small urban areas for which similar planning data variables had been utilized in the generation equations.

Internal trip-generation equations were reviewed for the urban areas of Pine Bluff, Arkansas (1965 population of 67,000); Fort Smith, Arkansas (1965 population of 80,000); and Kingsport, Tennessee (1966 population of 67,000). These 3 urban areas were used as the principal sources for the delineation of the initial trip-generation equations for the MUATS study.

In addition to the review of trip-generation equations developed for other small urban areas, a review also was made of the vehicle trip rate per capita and the percentage distribution between the trip purposes to be utilized in the MUAT study (Table 1, steps I and II). Table 2 gives the results of the review of the internal vehicle trips per capita for areas having a population of under 50,000. A value of 2.30 vehicle trips per capita was selected for use in developing the initial trip-generation equations for the MUAT study. Table 3 gives the results of the review of the percentage distribution of trips by trip purpose. The following trip-purpose distribution was selected for use in developing the initial trip-generation equations:

Trip Purpose	Percentage of Internal Trips
Home-based work	18
Home-based nonwork	44
Non-home-based	38
Total	100

TABLE 3

INTERNAL VEHICLE TRIPS BY PURPOSE

Study Area	Home-Based Work (percent)	Home-Based Nonwork (percent)	Non-Home-Based[a] (percent)
Murray, Kentucky	14.9	48.1	37.0
Pine Bluff, Arkansas	13.5	48.6	37.9
Fort Smith, Arkansas	22.5	38.7	38.8
Kingsport, Tennessee	20.4	42.9	36.7
Greenville, South Carolina	16.9	39.6	43.5
Memphis, Tennessee	21.2	41.5	37.3
Pulaski, Arkansas	21.8	47.3	30.9
Avg	18.8	43.8	37.4

[a]Includes truck and taxi trips.

The trip-generation equations from the 3 urban areas of Pine Bluff, Fort Smith, and Kingsport were combined with the selected trip-generation rate based on data given in Table 1 to develop the initial trip-generation equations to be used in the MUATS travel synthesis. The trip-generation equations resulting from this analysis are as follows.

Production Equations

1. Home-based work = 0.42 (population)
2. Home-based nonwork = 1.02 (population)
3. Non-home-based = 0.34 (population) + 1.57 (employment)

Attraction Equations

1. Home-based work = 1.268 (employment)
2. Home-based nonwork = 0.40 (population) + 2.22 (total employment - industrial employment)
3. Non-home-based = 0.34 (population) + 1.57 (employment)

The following total trips by trip purpose result from application of these equations and check closely with the initial trip-purpose distribution and per capita trip rate:

Trip Purpose	Internal Trips Number	Internal Trips Percent	Vehicle Trips Per Person
Home-based work	7,653	18.2	0.42
Home-based nonwork	18,589	44.4	1.02
Non-home-based	15,675	37.4	0.86
Total	41,917	100.0	2.30

As given in Table 2, there are ranges of reasonable internal trip-generation rates for use in this study. The selected trip rate of 2.30 internal vehicle trips per capita was considered reasonable for the initial trial synthesis of the internal vehicle trips. It was anticipated that this initial trip rate and, consequently, the trip-generation equations would require modification to achieve adequate balance between data and actual vehicular travel information. The section of this paper dealing with validity checks discusses the adjustments made in the trip rate during overall model calibration and presents the final trip rate and corresponding trip-generation equations.

Internal and External Attractions

Because an external origin and destination survey had been conducted for the Madisonville urban area, standard regression techniques were used in developing the internal and external trip-generation equation. The particular analysis procedure used in this study is referred to as a stepwise regression procedure. In this procedure, a sequence of multiple linear regression equations are computed in a stepwise manner in which 1 variable is added to the regression equation at each step. The variable added is the one that will show the greatest increase in the predictive accuracy of the equation. The IBM 1130 statistical package—stepwise multiple linear regression program—was used in performing this analysis.

The multiple linear regression procedure is very well documented in statistics textbooks, and no attempt will be made here to explain the procedure used and the various analytical and statistical tests conducted. The procedures used were in line with those recommended by the Bureau of Public Roads (10). Table 4 gives the series of equations developed, the selected equation, and related evaluation statistics.

TRIP-DISTRIBUTION MODEL

The trip-distribution model technique used in the Madisonville Urban Area Transportation Study was the gravity model. Because the basic approach is well known and accepted, a detailed discussion of all underlying principles and techniques is not repeated in this report (see 9, 10, 11, and 12 for specific details). However, pertinent examples are given in the ensuing discussions to illustrate the results obtained in the Madisonville study.

The basic gravity model expression relates trip interchanges between 2 zones in terms of the total trips produced in the zone of production, the total trips destined to the zone of attraction, and measures of the spatial separation of the 2 zones. Spatial separation relates primarily to travel time information and involves the development of relative tri-distribution rates for stipulated increments of time between zones. The gravity model expression is stated normally in the following general form:

$$T_{i-j} = \frac{P_i\, A_j\, F_{i-j}}{\sum\limits_{x=1}^{n} A_x F_{i-x}}$$

where

T_{i-j} = trips produced in zone i and attracted to zone j;
F_{i-j} = relative trip-distribution rate reflecting the spatial separation between zones i and j;
P_i = total trips produced in zone i;
A_j = total trips attracted to zone j;
n = total number of zones;
F_{i-x} = relative rate for distributing trips between zones i and x, where x varies from 1 to n; and
A_x = total trips attracted to zone x, where x varies from 1 to n.

Internal Trips

Application of the gravity model for distributing the internal trips requires that the zonal productions and attractions and the relative trip-distribution rates, F, be known. The synthesis of the internal vehicle trip productions and attractions by purpose was presented earlier in the section on trip generation analysis.

In preparation for the distribution of internal trips and for the calibration of the internal and external model, the existing major street and highway networks was prepared. The street and highway network was prepared in accordance with the specifications outlined by the Bureau of Public Roads in 1964 (12). Speeds on each facility contained in the existing network were estimated on the basis of functional classification and relative location within the urban area. Intrazonal driving times were calculated by applying an estimated speed for the local street system to average distances

TABLE 4

INTERNAL AND EXTERNAL TRIP ATTRACTIONS

Generation Equation	Cases[a]	Multiple Correlation Coefficient R	Mean	Standard Error of Estimate	Coefficient of Variation[b]	Student t-Test				Intercept as Percentage of Mean[c]
						1st Variation	2nd Variation	3rd Variation	4th Variation	
1. INT-EXT ATT = 77.6381 + 2.6163 (employment)	50	0.95	393	206	52	21	—	—	—	20
2. [d]INT-EXT ATT = 44.7115 + 0.3434 (population) + 2.5928 (employment)	50	0.97	393	172	44	5	25	—	—	11
3. INT-EXT ATT = -27.1529 + 0.3154 (population) + 2.3491 (employment) + 0.9009 (public employment)	50	0.97	393	153	39	5	21	4	—	7
4. INT-EXT ATT = -3.9865 + 0.2859 (population) + 1.1695 (employment) + 1.2963 (commercial employment) + 2.0530 (public employment)	50	0.98	393	139	35	5	3	3	5	1

[a]The number of zones used as observations in developing the equation.
[b]The standard error of estimate divided by the mean and expressed as a percentage.
[c]The constant term in the equation.
[d]Selected equation for forecasting purposes.

measured from the centroid of each zone to its extremities. These intrazonal times averaged 1 min in the Madisonville area. Terminal times were estimated based on consideration of the time spent walking to and from a parking place and the time spent looking for a parking place. The resulting terminal times for the MUAT study varied from 1 min in residential zones to 3 min in the CBD. Intrazonal travel time was then computed by adding intrazonal driving time to twice the terminal time.

Development of the initial travel time factors to be used in the MUAT study was based on a review of similar information of other small urban areas. These urban areas included Fort Smith, Arkansas; Pine Bluff, Arkansas; Kingsport, Tennessee; Murray, Kentucky; and average data from 7 cities in Iowa. The resulting plots of this information are shown in Figure 1 for each of the 3 trip purposes used in the internal trip-generation analysis previously discussed. In some instances, various curves were shifted so that all curves would pass through a common point. However, the basic shape of the curve and, hence, the relative values represented by the curve were not altered by making the shift. This common point was selected at F = 100 and t = 4 min, inasmuch as travel time factors for the first 3 to 4 min are of doubtful value in a gravity model distribution. Table 5 gives the initial travel times selected for use in the distribution of internal trips for the Madisonville study and based on the values obtained from the analysis of

TABLE 5

INITIAL TRAVEL TIME FACTORS BY TRIP PURPOSE

Travel Time[a] (min)	Work	Nonwork	Non-Home-Based
1	200	360	310
2	156	250	200
3	124	160	135
4	100	100	100
5	85	75	72
6	72	60	53
7	62	45	41
8	54	36	32
9	47	27	25
10	41	19	21
11	36	14	18
12	33	11	16
13	29	9	13
14	26	7	11
15	22	6	10
16	20	5	8
17	16	4	7
18	14	3	6
19	11	2	5
20	9	1	4
21	7	—	3
22	5	—	3
23	3	—	2
24	2	—	1
25	1	—	—

[a]Includes terminal time.

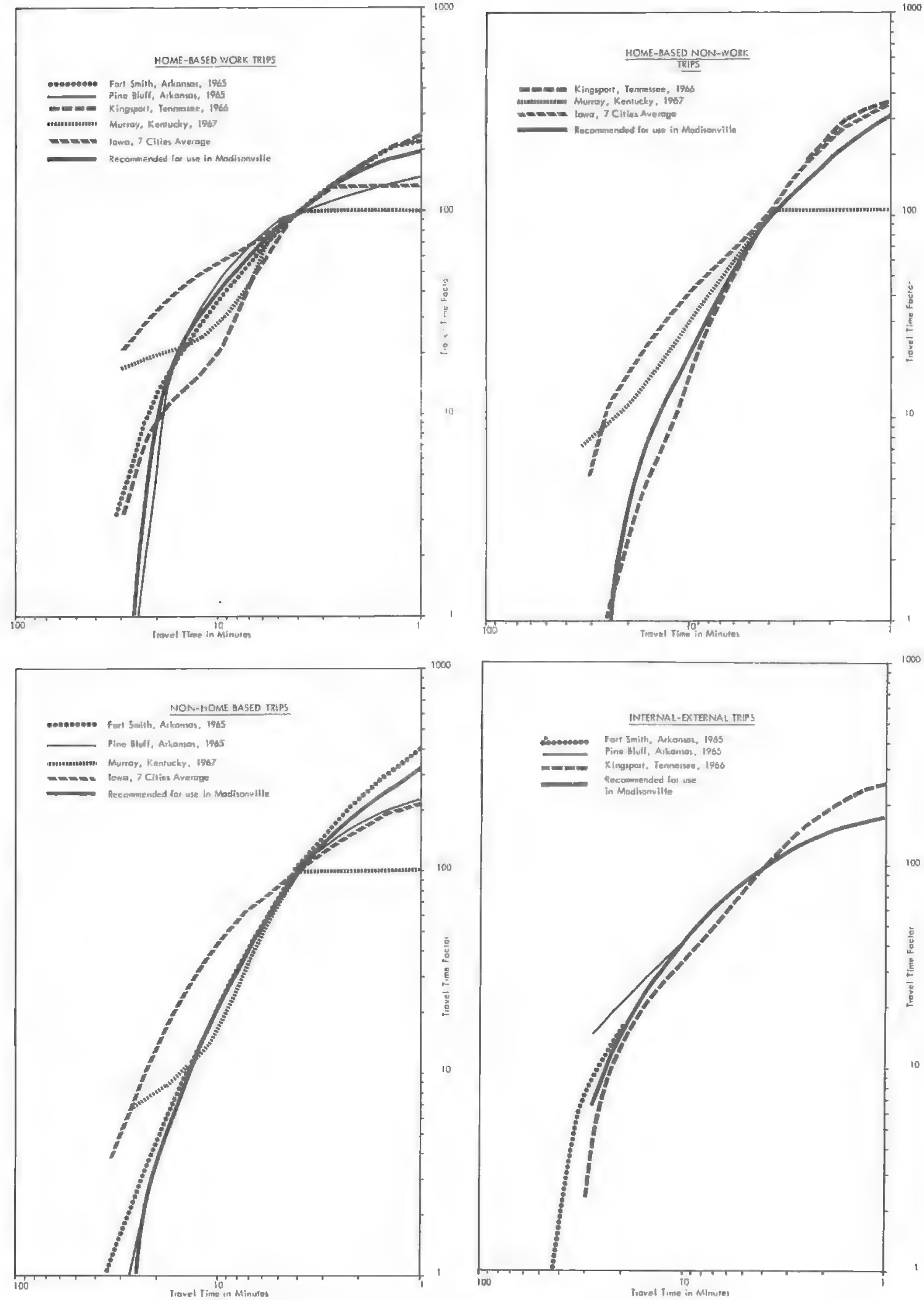


Figure 1. Travel time factor comparisons.

other small urban areas shown in Figure 1. The trip length frequencies shown in Figure 2 are derived from application of the gravity model expression to the zonal productions and attractions for each purpose obtained from the generation analysis.

As shown in Figure 2, the trip-length frequency for each of the 3 trip purposes used in the study was altered between the first and third trial calibrations performed on the models during the process of obtaining an adequate comparison between the model and the actual ground information within the study area. The details of the changes made between each trial calibration are presented in a later section of this report dealing with the validity checks performed on the travel models.

Internal and External Attractions

Development of the initial travel time factors for the internal and external trips also was based on a review of information from the same urban areas considered for the

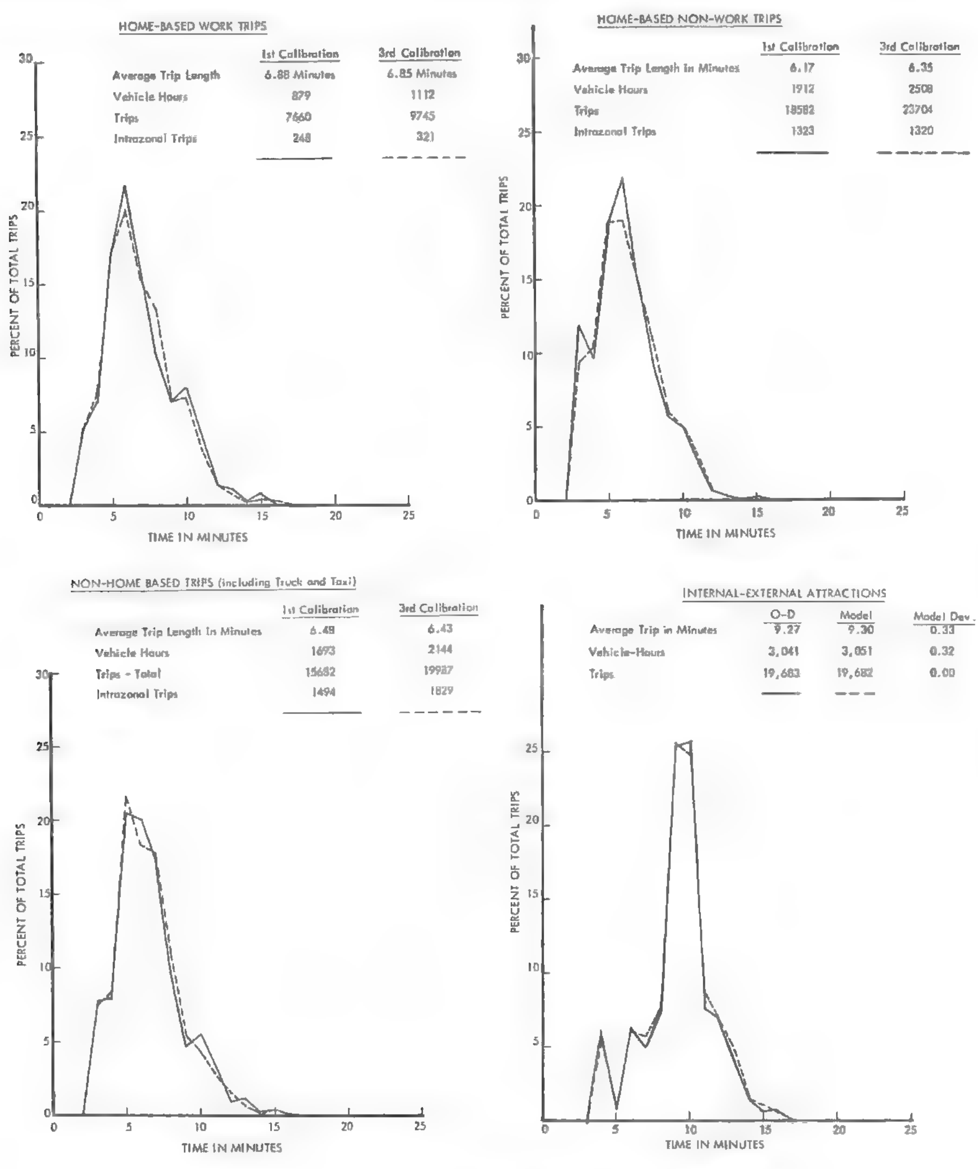


Figure 2. Trip length frequency distributions.

TABLE 6

RECOMMENDED TRAVEL TIME FACTORS BY TRIP PURPOSE

Travel Time[a] (min)	Work	Nonwork	Non-Home-Based	Internal and External
1	200	250	220	170
2	156	170	160	150
3	124	130	125	120
4	100	100	100	100
5	85	75	72	86
6	72	60	53	76
7	62	50	41	64
8	54	41	32	56
9	47	34	25	48
10	41	28	21	42
11	36	24	18	38
12	33	21	16	34
13	29	18	13	31
14	26	16	11	28
15	22	14	10	25
16	20	12	8	23
17	16	11	7	21
18	14	10	6	19
19	11	8	5	17
20	9	6	4	15
21	7	4	3	14
22	5	2	3	12
23	3	1	2	11
24	2	—	1	10
25	1	—	—	9

Note: Includes terminal time.

internal distribution model. The resulting travel time factors for the internal and external trip distribution selected for use in the initial trial calibration of the gravity model are given in Table 6 and are shown in Figure 1.

An initial gravity model calibration was made on the trip data for the Madisonville study to obtain trip-length frequency information for comparison purposes. Figure 2 shows the comparison of the trip-length frequency of the external origin and destination trips and the trip-length frequency obtained from the first trial calibration of the gravity model. Figure 2 also shows comparisons of the origin and destination data and model data for the percentage of deviation in trips, vehicle-hours, and average trip length. The model deviation is well within the 3 percent value generally considered the maximum acceptable percentage of deviation. Thus, the values given in Table 6 were selected as the final travel time factors for internal and external trip distribution.

VALIDITY CHECKS PERFORMED ON THE TRAFFIC MODELS

After the internal trips were distributed and the internal and external gravity model was calibrated, various analytical and statistical tests were applied to verify the results of the mathematical models. The basic procedure utilized is given in Table 1 (step V). In brief, the external trip table (internal and external and through trips) and the total trip table (internal, internal and external, and through trips) were assigned

separately to the existing major street and highway networks. Vehicle-mile checks and screenline comparisons between the model-assignment values and the ground counts served as the major overall validity checks on model adequacy. The standard "all-or-nothing" assignment procedure and the package of programs developed for the IBM 1130 computer were used. The results of these assignments were compared to the ground counts on key links throughout the network.

To obtain the vehicle-mile validity check mentioned previously required ground-count information for all links in the existing major street and highway network. Inasmuch as blanket coverage of ground-count information was not available, volumes on various links throughout the existing network were estimated. This estimation was based on expansion of historical count information, expansion of peak-hour turning volume counts at the signalized intersections, and current count information for street segments having characteristics and use similar to those of the segments being estimated.

Because the ground count volume on a substantial number of links within the network was estimated, a check was performed to determine the overall adequacy of using these estimated volumes. This check consisted of a vehicle-mile comparison of the assignment volumes and the ground count volumes for all links and for selected links within the network (links where actual ground-count information was available were termed "selected links"). The results of this comparison are shown in the following:

Links	Vehicle-Miles: Ground Counts	Vehicle-Miles: Assignment (third calibration)	Percent Deviation
Selected	137,600	131,700	4.3
All	213,800	204,700	4.2

It was concluded that the estimated ground counts were reasonable on an overall basis.

Based on the preceding method of comparing the model assignment volumes to the ground count volumes on those links crossing each of the selected screenlines and of comparing the total model assignment vehicle-miles to the ground count vehicle-miles, the initial assignment to the network indicated that the mathematical models were under-predicting travel within the study area. The results of this comparison for the first calibration are given in Table 7.

The internal and external trips and the through trips contained in the model assignment values given in Table 7 were obtained from actual roadside interviews in the external origin and destination survey conducted for the study area. Thus, for the model development it was accepted that these trips reproduce the corresponding ground counts to the required degree of accuracy and, therefore, require no adjustment. Consequently, to obtain a closer check between the model assignment and the ground count values required that the internal trip rate be increased by about 28 percent prior to reassignment of the total trip matrix.

In addition to increasing the internal trip-generation rate from the initial value of 2.30 trips per capita to a new value of 2.93 trips per capita, the second calibration of the model also incorporated speed adjustments on selected links within the network, particularly in and around the CBD so that a closer comparison could be obtained between the model assignment and the ground counts. After these adjustments were made, the internal trips were redistributed by using the initial travel time factors given in Table 5; and the resulting total trip matrix (internal, internal and external, and through trips) were reassigned on an all-or-nothing basis to the revised network.

The screenline and vehicle-mile comparisons mentioned previously were made on the results of the second calibration traffic assignment. The results of these comparisons are also given in Table 7. Overall, the model vehicle-miles were about 8 percent less than the corresponding vehicle-miles computed from the ground counts. The screenline analysis also indicated that the models were still underestimating trips by about 8 percent. However, the results of the screenline analysis for the second calibration traffic assignment showed that the volumes crossing a number of screenline segments were being over-predicted by the model whereas other segments were being

TABLE 7

LOADED NETWORK SCREENLINE COMPARISONS

Screenline	Section	Ground Count	Model	
			Calibration 1	Calibration 2
2	Cordon-B	7,000	6,545	7,356
	B-C	32,939	19,813	24,987
	C-D	4,304	5,364	5,869
	Total	44,243	31,722	38,212
B	Cordon-2	12,000	9,804	11,298
	2-cordon	7,220	8,017	8,262
	Total	19,220	17,821	19,560
C	Cordon-2	5,837	4,931	5,066
	2-cordon	33,066	27,795	30,753
	Total	38,903	32,726	35,819
D	Cordon-2	1,610	1,666	1,694
	2-cordon	24,357	20,447	22,090
	Total	25,967	22,113	23,784
Total		128,333	104,382	117,375
Total vehicle-miles		213,769	189,906	202,860

Note: See Figure 3 for screenline locations.

underpredicted. Thus, it was concluded that adjustment of the initial friction factors used in distributing the internal trips was probably required to obtain a more even distribution.

Several other factors also pointed to the adjustment of the travel time factors as being the most desirable adjustment prior to the third calibration. These factors consisted of the magnitude of the internal trip-generation rate and the problem of possible double crossings of the screenline segment traversing the central area. Because the internal trip rate of 2.93 vehicle trips per capita resulting from the second calibration was approaching an upper limit for urban areas of relatively small size, further adjustment of this internal trip rate was considered undesirable until such time as the distribution of trips was deemed accurate.

Moreover, the largest single difference between the ground-count volume and the model-assignment volume crossing a given screenline segment occurred on screenline 2 from section B to C (i.e., the portion of screenline 2 traversing the CBD). Figure 3 shows the screenline locations. Because of the location of this screenline segment and the high peak-hour parking occupancies noted for on-street parking facilities within the CBD, the possibility existed that the ground counts in this area reflected a considerable number of double crossings resulting from circulating traffic. Consequently, adjustment of the internal trip distribution was considered more appropriate for the third calibration than an additional adjustment in the internal trip-generation rate.

Because the major discrepancies between the ground-count volumes and the model-assignment volumes on a link-by-link basis occurred in and around the central area, the travel time factors for the medium-length trips were selected for adjustment. This adjustment was further limited primarily to the nonwork trips because of the heavy commercial orientation of the central area. The travel time factors resulting from this revision are given in Table 6. These revised travel time factors were used to redistribute and reassign the trip productions and attractions on an all-or-nothing basis.

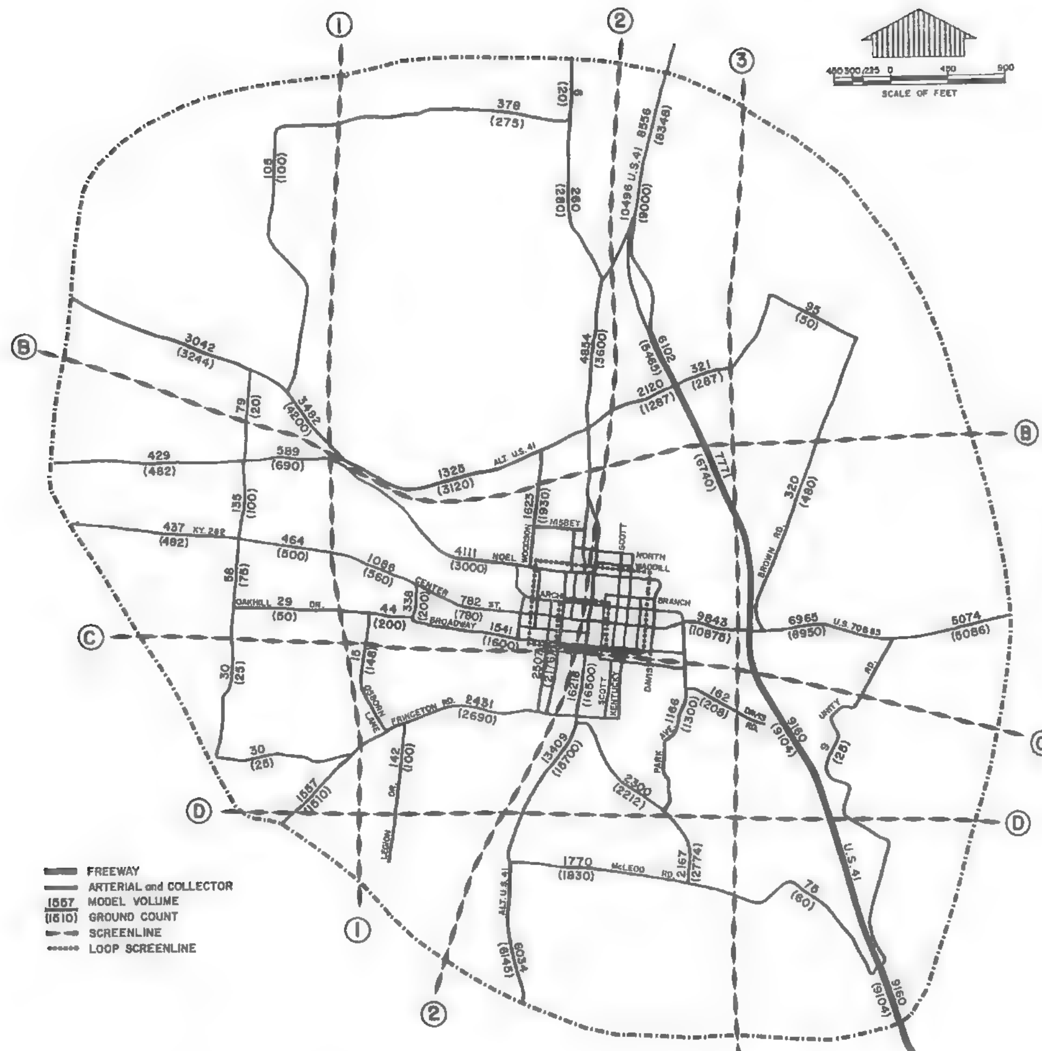


Figure 3. Screenline comparisons of model-assigned trips versus ground counts.

The results of the third calibration traffic assignment are given in Table 8. Analysis of Table 9 reveals that the models reproduced the ground counts within about 5 percent accuracy on an area-wide basis for both the screenline and the vehicle-mile comparison. Figure 3 shows the locations of the screenlines as well as the assigned volumes and the ground counts on selected links throughout the system. Thus, these results indicate that the third calibration of the mathematical model reproduced the existing travel pattern with acceptable accuracy.

This conclusion was verified by selecting a number of additional screenlines. The location of these screenlines is shown in Figure 3. These additional screenlines consisted of 2 north-south and 2 loop screenlines. The location of the north-south screenlines (i.e., screenlines 1 and 3) was selected to bracket the well-developed portion of the existing urban area. The location of the 2 loop screenlines was selected to evaluate the portion of the internal trips traveling to and from the CBD and to minimize the

TABLE 8

LOADED NETWORK SCREENLINE COMPARISON—CALIBRATION 3

Screenline	Section	Ground Count	Assigned Volume
1	Cordon-B	4,300	3,587
	B-C	1,240	1,082
	C-D	2,250	2,312
	Total	7,790	6,981
2	Cordon-B	7,000	7,436
	B-C	32,939	29,748
	C-D	4,304	2,300
	Total	44,243	39,484
3	Cordon-B	287	321
	B-C	17,030	17,609
	C-D	208	162
	D-cordon	980	2,317
	Total	18,505	20,409
B	Cordon-1	20	79
	1-2	11,980	11,582
	2-3	6,740	7,771
	3-cordon	480	320
	Total	19,220	19,752
C	Cordon-1	25	30
	1-2	5,812	4,698
	2-3	23,586	21,396
	3-cordon	9,480	9,169
	Total	38,903	35,293
D	Cordon-1	1,510	1,557
	1-2	100	142
	2-3	15,228	13,428
	3-cordon	9,129	9,169
	Total	25,967	24,296
Total		154,628	146,215
Inner loop		71,323	63,894
Outer loop		67,394	70,941
Total vehicle-miles		213,769	204,687

Note: See Figure 3 for screenline locations.

potential problem of double screenline crossings previously discussed. The results of these screenline comparisons for the third calibration are also given in Table 8.

As an additional test on the adequacy of the calibrated model, the assigned volumes and the ground counts were compared by using a standard Bureau of Public Roads program (modified for use on the IBM 1130 computer) that subdivides the individual links into groups based on the ground-count traffic volume. For each volume group, the computer program tabulates the number of links in the volume group, the average volume for each group, the average difference between the ground counts and the model volumes, and the root mean square error of the differences between the ground counts and the model trips. These statistics for each volume group are given in Table 9.

TABLE 9

LOADED NETWORK COMPARISON BY LINK ONE-WAY VOLUMES

Volume Group	Number of Links	Percentage of Total Links	Average Ground Count	Average Difference	Root Mean Square Error	Percentage of RMSE
Less than 1,999	359	72.3	714	-64	570	80
2,000-3,999	78	16.0	2,790	+291	1,071	38
4,000-5,999	42	8.4	4,910	+690	1,079	22
6,000-7,999	10	2.1	7,050	+914	1,051	15
Over 8,000	6	1.2	8,520	+555	870	10
Total	495	100.0	1,620	+83	750	46

Calculation of the percentage of root mean square error does not provide an absolute check on the adequacy of the calibrated model because of the lack, for the Madisonville study, of an internal origin and destination sampling rate. However, this calculation can provide an indication of whether the calibrated movel is generally adequate or generally inadequate. As given in Table 9 the root mean square error (percentage of RMSE) is reasonable for volume groups of more than 2,000 vehicles per day. The relatively large percentage of root mean square error for those links having an average daily traffic volume of less than 2,000 vehicles per day is relatively insignificant when related to the service volume of a 2-lane highway. That is, from a planning point of view, the errors indicated by this large root mean square error do not change the lane requirements on a specific facility carrying 2,000 vehicles per day or less.

Moreover, this relatively large discrepancy is attributable to the problems inherent in the basic traffic assignment procedure as well as the accuracy of the ground counts used for comparison purposes in this calculation. Because of the small size of the urban area, travel paths differing by fractions of a minute or a second may result and the rounding by the computer may be biased. Because the entire existing major street and highway network was used in the model calibration, the likelihood of large discrepancies between the ground-count volume and the assignment volume is relatively high in those instances of paralleling facilities and of network links located adjacent to zone centroid connectors. In addition, many of the ground-count volumes on these low-volume links were estimated for the purposes of comparison and, therefore, may contain considerable error when analyzed on an individual link basis. However, as indicated by the values given in Table 9, the majority of the volume groups have a generally acceptable root mean square error.

Based on the preceding analyses, the trip-generation and trip-distribution models resulting from the third calibration traffic assignment are recommended for the distribution of future trips. The following are the recommended trip-generation equations:

Production Equations

1. Home-based work = 0.5355 (population)
2. Home-based nonwork = 1.3005 (population)
3. Non-home-based = 0.4265 (population) + 2.0018 (employment)

Attraction Equations

1. Home-based work = 1.6167 (employment)
2. Home-based nonwork = 0.5100 (population) + 2.8305 (total employment - industrial employment)
3. Non-home-based = 0.4265 (population) + 2.0018 (employment)
4. Internal-external = -44.7115 + 0.3434 (population) + 2.5928 (employment)

Table 6 gives the recommended travel time factors by trip purpose resulting from the final calibration of the gravity model. The recommended trip-length frequency distributions for internal trips and internal and external trips are shown in Figure 2 (calibration 3).

CONCLUSIONS

1. The developed model gave good results and is recommended for forecasting future trips. Forecasts based on the trip-generation equation will probably be on the low side. The trip-generation equations have population and employment as the independent variables. Population and employment have been growing and are expected to continue growing at a slower rate than vehicle registration and at an even slower rate than vehicular trips. Control totals for each trip purpose should be established independently of the trip-generation equations.

2. The overall cost of this phase of the Madisonville study is about $6,000. This paper demonstrates that internal trip patterns can be economically synthesized in small urban areas. However, more research is needed to determine the trip-length frequency distribution, factors affecting them, and influence of movement of a big employer into or out of the area or from 1 side of the area to the other.

3. More research is needed to determine the magnitude of intraurban area vehicle trips made by out-of-area residents. In small urban areas, the central city is the focal point of surrounding rural areas. These intraurban area trips, commonly referred to as nonreported trips, could be collected at a low additional cost by including extra questions to this effect on the external cordon survey interview form.

4. The model has the advantages of being simple and economically feasible. It uses 3 socioeconomic factors that can be easily forecasted—population, total employment, and industrial employment. The general procedure is applicable to any small urban area. However, the exact model needs to be developed and calibrated for each small urban area separately.

REFERENCES

1. Mathematical Traffic Model for the Madisonville Urban Area Transportation Study. Harland Bartholomew and Associates, Tech. Memo. A, 1969.
2. Mathematical Traffic Model for the Fort Smith Urban Area Transportation Study. Harland Bartholomew and Associates, Tech. Memo., 1966.
3. Mathematical Traffic Model for the Pine Bluff Urban Area Transportation Study. Harland Bartholomew and Associates, Tech. Memo., 1967
4. Mathematical Traffic Model for the Kingsport Urban Area Transportation Study. Harland Bartholomew and Associates, Tech. Memo., 1968.
5. Harmelink, M. D., Harper, G. C., and Edwards, H. M. Trip Production and Attraction Characteristics in Small Cities. Highway Research Record 205, 1967, pp. 1-19.
6. Jefferies, W. R., and Carter, E. C. Simplified Techniques for Developing Transportation Plans: Trip Generation in Small Urban Areas. Highway Research Record 240, 1968, pp. 66-87.
7. Better Transportation for Your City. National Committee on Urban Transportation, Chicago, 1958.
8. Shuldiner, P. W. An Analysis of Urban Travel Demands. Northwestern Univ. Press, 1962.
9. Calibrating and Testing a Gravity Model for Any Size Urban Area. U.S. Bureau of Public Roads, 1965.
10. Guidelines for Trip Generation Analysis. U.S. Bureau of Public Roads, June 1967, 118 pp.
11. Summary of Urban Travel Data. U.S. Bureau of Public Roads, 1967.
12. Traffic Assignment Manual. U.S. Bureau of Public Roads, 1964.
13. Whitmore, E. R. Graphical and Mathematical Investigation of the Differences in Traveltime Factors for the Gravity Model Trip Distribution Formula in Several Specific Areas. Univ. of Tennessee, 1965.

A MULTIPATH TRAFFIC ASSIGNMENT MODEL

Robert B. Dial, Alan M. Voorhees and Associates, Inc., McLean, Virginia

This paper describes the mechanics of a novel traffic assignment model that is able to assign coterminous trips to alternative routes without resorting to reiteration or path enumeration. Under a particular definition of "reasonable path," the model satisfies certain common-sense requirements. The model is a 2-pass Markov model. It calculates node and link transition probabilities in 1 examination of the network and assigns trips in a second examination, when it diverts trips entering a node to all reasonable links ending at the node. The model never explicitly examines a path, and it is not in any sense an "optimization model." It assigns trips to all reasonable paths simultaneously in such a way that the resulting effect is identical to what would have been obtained had each path been assigned trips separately under certain choice probability assumptions. Thus, compared to other multipath assignment techniques, the model is theoretically attractive and computationally very efficient. Presented also are 2 algorithms that differ in their definition of a reasonable path and in the number of times each is executed to assign all trips from a given origin node.

•IN MOST widely used traffic assignment models, all trips between a fixed origin and destination are assigned to the links constituting a single shortest connecting path. (In this article, link "time" and "length" are used interchangeably to mean neither's literal definition. Here they are names for the link disutility measure representing travel cost or impedance. We assume a link's disutility is always a positive number. A path's length is the sum of the disutility of the links that constitute it. The shortest path is one whose links sum to the smallest total disutility—whatever it may be.) This latter technique has been designated the "all-or-nothing" assignment. Because of the effects of trip volumes on travel time and the trip-maker's nondeterministic choice function on route selection, all-or-nothing assignment is known to contradict actual trip behavior; and the link-volume output of these traffic assignment models is sometimes inaccurate to the point of compromising the transportation planner's design decisions.

Many transportation planners feel that a traffic assignment model would be much more useful if it could efficiently reflect, to some degree, the nonoptimal behavior of trip-makers. The quality of the planner's decisions could be improved, and the cost of arriving at them could be decreased. A highway system, particularly when operating at near-capacity volumes, provides many alternate paths that vary slightly with respect to length between the same origin and destination. A realistic model would be a "multipath" assignment model, which would apportion trips to all of these paths in a probabilistic manner reflecting each path's relative likelihood of use.

An easy 3-step solution to this problem would be a model that would (a) relate path choice to path characteristics; (b) find all paths between a given origin and destination; and (c) using relationships found in step a, apportion trips to the paths on the basis of their characteristics. Taken literally, the preceding method has little utility. Even though such a model is fairly easy to design and implement, the large size (4,000 to 15,000 nodes) of the networks precludes this obvious solution. There are too many paths. Computers are not yet fast enough to perform the implied computation in a

Sponsored by Committee on Transportation Forecasting.

reasonable amount of time and, thus, render such a model uneconomical to use. The economics of model utilization demand that the differential value of a sophisticated multipath traffic assignment model be greater than the differential value of the output minus the differential cost of obtaining it. The cost of the output includes data preparation and computation cost. As yet, no "pure" multipath assignment model has managed to achieve a positive differential value. The computation time required to search out and evaluate alternative paths costs more than the information is worth. The proof of the statement is implied by the technique's nonuse.

The probablistic assignment model presented in the following is an attempt to circumvent the path enumeration problem. It assigns trips to all "reasonable" paths simultaneously in such a way that the resulting effect is identical to what would have been obtained had each path been assigned trips separately under certain choice probability assumptions. Compared to other multipath techniques, the model is theoretically attractive and computationally very efficient. On the theoretical side it displays some highly desirable characteristics seldom present in other techniques.

Under a particular definition of "reasonable path," theoretical appeal comes from the model satisfying the 3 following functional specifications:

1. The model gives all reasonable paths between a given origin and destination a nonzero probability of use whereas all unreasonable paths have a probability of zero;
2. All reasonable paths of equal length have an equal probability of use; and
3. When there are 2 or more reasonable paths of unequal length, the shorter has the higher probability of use.

Computational efficiency and flexibility result from the model satisfying 2 additional functional specifications:

4. The model does not explicitly enumerate the paths it loads, but all reasonable paths between a given origin and destination are loaded simultaneously; and
5. The user is able to control the path diversion probability by assigning a value to a parameter θ that affects the slope of the "diversion curve."

Computationally, the model can be called a 2-pass Markov model. It calculates node and link transition probabilities during 1 look at the network and assigns trips during a second look when it diverts trips entering a node to all reasonable (efficient) links ending at the node. In this way, it assigns trips simultaneously to an entire set of reasonable paths. The model never explicitly examines a path, and it is not in any sense an "optimization model."

In the next 2 chapters, the mechanics of 2 models are presented. Both of them satisfy the 5 preceding specifications. Both are Markov models that probabilistically divert trips from nodes to competing converging links. The 2 models differ in their definition of an efficient path and in the number of times their implementing algorithm is executed to assign all trips from a given origin node. This article is quite informal. For a lengthier and more rigorous discussion, the reader is referred to another paper by Dial (70), which presents algorithms in more detail, describes their computer implementation, and provides complete formal justification for some of the unsupported claims made in the following sections.

MODEL 1—PROBABILISTIC MULTIPATH ASSIGNMENT

This first model requires that a reasonable path between nodes o and d be an efficient path, composed only of links possessing the 2 following properties:

1. The initial node of the link is closer to the origin node o than is its final node, and
2. The final node of the link is closer to the destination node d than is its initial node.

These dual constraints restrict the set of efficient paths to those relating symmetrically to the origin and destination nodes. This duality requires that the assignment algorithm be executed once for each pair of nodes o and d. Although this is an understandable

requirement, it is time consuming in execution. First, the shortest path length from each node to d must be known, and, second, there are many o-d pairs. The algorithm is described in the following section.

Algorithm 1

Preliminaries—To assign y trips between origin node o and a destination node d requires that the 4 following items be known for each node i: $p(i)$ = the shortest path distance from o to i; $q(i)$ = the shortest path distance from i to d; I_i = the set of all links whose initial node is node i; and F_i = the set of all links whose final node is node i. Letting link $e = (i, j)$ have length $t(i, j)$, we can calculate for each link e its likelihood:

$$a(e) = \begin{cases} \exp\theta\,[p(j) - p(i) - t(i,j)] & \text{if } p(i) < p(j) \text{ and } q(j) < q(i) \\ 0 & \text{if otherwise} \end{cases}$$

Having thus defined a(e), we can describe the algorithm as a 2-pass process, which need concern itself only with those links whose a(e) is not zero.

Forward Pass—By examining all nodes i in ascending sequence with respect to p(i), their distance from the origin, we can calculate for each link e in I_i its link weight:

$$w(e) = \begin{cases} a(e) & \text{if } i = o \text{ (the origin node)} \\ a(e) \sum_{e' \text{ in } F_i} w(e') & \text{if otherwise} \end{cases}$$

When the destination node d is reached, the next step is undertaken.

Backward Pass—Starting with the destination node d, we can examine all nodes j in descending sequence with respect to p(j). A trip volume x(e) is assigned to each link e in F_j as follows:

$$x(e) = \begin{cases} y \cdot w(e) \Big/ \sum_{e' \text{ in } F_j} w(e') & \text{if } j = d \text{ (the destination node)} \\ w(e) \sum_{e' \text{ in } I_j} x(e') \Big/ \sum_{e' \text{ in } F_j} w(e') & \text{if otherwise} \end{cases}$$

When the origin node o is reached, the assignment is complete. Notice that, in the algorithm, the funxtions x and w are defined recursively, and thus the order of their calculation must be as specified.

Example—In the simple grid network shown in Figure 1, we assume that all link times are 2, except on the links forming the bisecting horizontal path between node 11 and node 15, where the link times are all unity. We assign 700 cars making the trip from node 1 to node 25 as follows:

1. All pertinent preliminary data are shown in Figure 1: Above and to the left of each node is p(i), the distance the node is from node 1; below and to the right of each node is q(i), the distance the node is from node 25; the exiting link sets I_i; and the entering link sets F_i. Assume that the parameter θ is unity in the definition of a(e). Then the nature of the network allows only 2 values for the exponent of a(e): 0 or -1. Therefore, a(e) is either 0, 1, or 0.3679. The appropriate value for a(e) is posted above those links for which it is nonzero. Where a(e) is zero, no arrows or values appear on the links. Arrowless links will receive no trips because a(e) and, therefore, w(e) are zero.

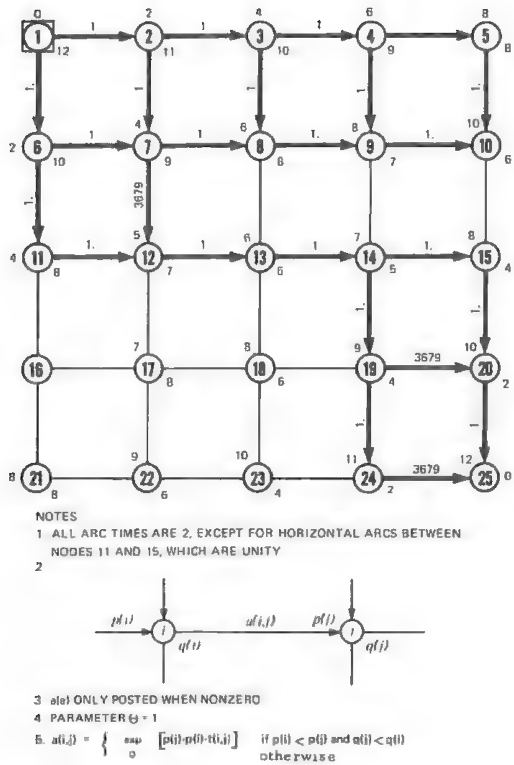


Figure 1. Preliminary data for Algorithm 1.

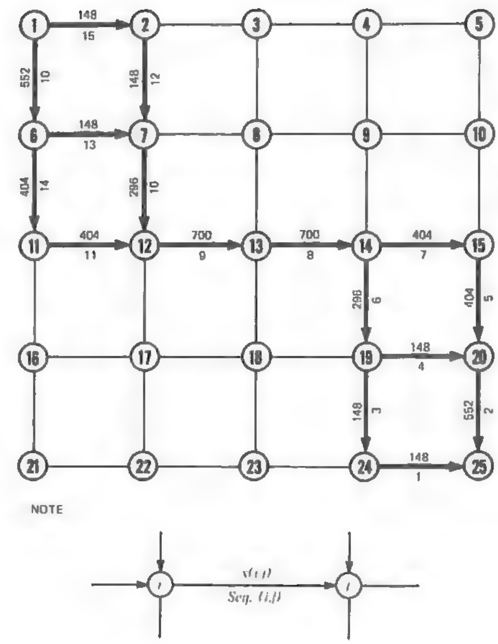


Figure 3. Results of backward pass.

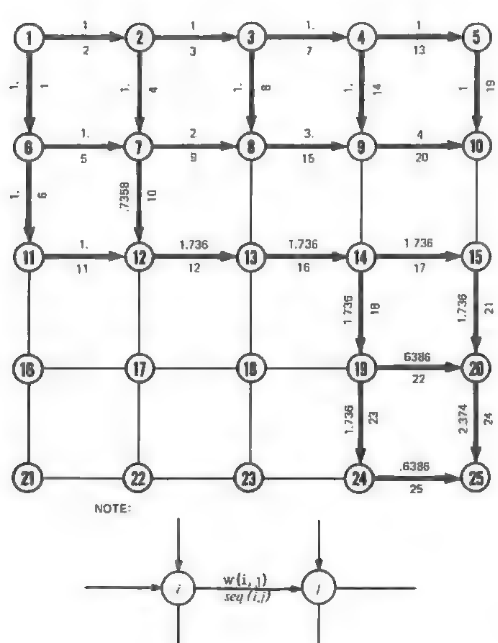


Figure 2. Results of forward pass.

2. Forward pass results are shown in Figure 2. a(e) has been replaced with w(e). The integer below each link indicates the sequence in which the links were processed. For example, link (20,25) was the 24th link processed, and w(20,25) = a(20,25) [w(15,20) + w(19,20)] = 1(1.736 + 0.6386) ≅ 2.374.

3. Backward pass results are shown in Figure 3. w(e) has been replaced with x(e), the link volume, and the sequenced numbers have been changed to correspond to the links' computational sequence in the backward pass. For example, link (7,12) was the 10th link processed and x(7,12) = w(7,12) · x(12,13)/[w(7,12) + w(11,12)] = 0.7358 · 700/(0.7358 + 1) ≅ 296.

Figure 3 shows the symmetry of the assignment around the dominating freeway axis, which reflects the symmetry of the network. There are 9 efficient paths between nodes 1 and 25. A unique shortest path has a length of 12 units. Four paths are 13 units long. Four have a length of 14. All 9 paths have been simultaneously assigned trips in a single execution of the algorithm.

Figure 4 shows the volume effectively assigned to each of these 9 paths. Each subfigure depicts an efficient path composed of the links indicated by heavy arrows. Posted beneath its graph are the length of the path and the number of trips effectively assigned to it. For example, one finds the shortest path between nodes 1 and 25. Its length is 12 and the number of trips the algorithm effectively assigned to it is equal to 232. Figure 4 shows one of 4 efficient paths between nodes 1 and 25 whose length is 13. The inferred trip volume on this path is 85. Figure 4 shows a 14-unit long path that accommodates 32 trips.

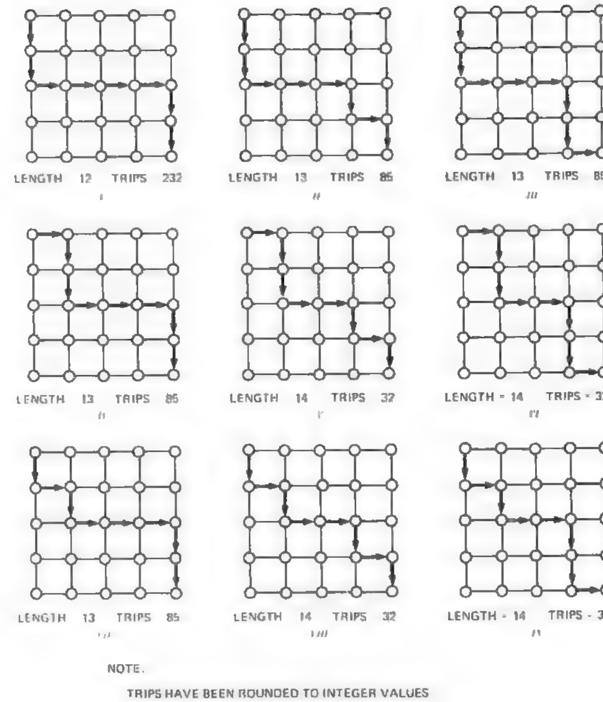


Figure 4. Efficient paths and volumes.

Justification of Algorithm 1

As mentioned previously, detailed justification of the assignment algorithm exists elsewhere. Here only a sketchy proof is given that the algorithm does in fact satisfy the 5 functional specifications given at the beginning of this paper.

In the preliminaries of the algorithm, we define the likelihood of a link e = (i,j) as

$$a(e) = \begin{cases} \exp \theta\, [p(j) - p(i) - t(i,j)] & \text{if } p(i) < p(j) \text{ and } q(j) < q(i) \\ \\ 0 & \text{if otherwise} \end{cases} \tag{1}$$

Notice that the exponent is directly proportional to p(j) - [p(i) + t(i,j)], which is the nonpositive difference between the shortest distance to node j and the length of the shortest path to node j that uses link (i,j). Roughly speaking, a(e) is a kind of "shadow cost" of using link e.

It is assumed that the probability of using a particular (simple) path P is directly proportional to the product of the likelihood of the links in the path; that is,

$$\text{prob}(P) = k \prod_{e \text{ in } P} a(e) \tag{2}$$

Thus, prob(P) is nonzero if and only if the path P is efficient. This verifies specification 1.

By substituting Eq. 1 into Eq. 2, the probability of an efficient path can be written as

$$\text{prob}(P) = k \prod_{e = (i,j) \text{ in } P} \exp \theta\, [p(j) - p(i) - t(i,j)] \tag{3}$$

$$\text{prob}(P) = k \exp \theta \sum_{(i,j) \text{ in } P} [p(j) - p(i) - t(i,j)] \tag{4}$$

$$\text{prob}(P) = k \exp \theta \left[p(d) - \sum_{(i,j) \text{ in } P} t(i,j) \right] \tag{5}$$

The transition from Eq. 4 to Eq. 5 follows from the fact that consecutive links in a path share a node; and, for any given node $i \neq o$, d on the path P, p(i) will appear in the summation in Eq. 4 exactly twice and with opposite sign.

Because p(d) is the shortest path distance from o to d, and θ is a positive constant, Eq. 5 shows that the model satisfied specifications 2 and 3. The value of the summation in Eq. 5 is just the length of path P. Hence, the exponent in Eq. 5 is nonpositive and becomes more negative as the length of P increases. That is to say, prob(P) decreases with increasing path length.

To show that specification 5 is satisfied, we only have to show that the calculated link volumes are obtained in a manner consistent with Eq. 2, because the algorithm obviously does not enumerate paths. This is shown by proving that the algorithm diverts trips from each node according to appropriate conditional link probabilities. A conditional link probability is the probability that a trip between o and d will use a particular line $e = (i,j)$, given that it goes through the link's final node. This probability can be formally stated as

$$\text{prob}[(i,j)|j] = \text{prob}[(i,j), j] / \text{prob}(j) \tag{6}$$

$$\text{prob}[(i,j)|j] = \text{prob}[(i,j)] / \text{prob}(j) \tag{7}$$

$$\text{prob}[(i,j)|j] = \text{prob}[(i,j)] \Big/ \sum_i \text{prob}[(i,j)] \tag{8}$$

Although it is obvious that the probability of using a link (i,j) is just

$$\text{prob}[(i,j)] = \sum_{\{P:\,(i,j) \text{ in } P\}} \text{prob}(P) \tag{9}$$

It is useful to write Eq. 9 in a more elaborate form, to facilitate cancellation of common factors in the numerator and denominator of Eq. 8. To this end, notice that an efficient path through link (i,j) can be partitioned into 3 sets of links: (a) P_i = {all links topologically preceding link (i,j)}; (b) {(i,j)}; and (c) P_j = {all links topologically following link (i,j)}. Now if $\mathbb{P}_i$ is the family of all P_i representing partition 1 of any efficient path from o to d, and $\mathbb{P}_j$ is the family of partition 3 of all efficient paths between o and d, then

$$\sum_{\{P:\,(i,j) \text{ in } P\}} \text{prob}(P) = k\, a(i,j) \left[\sum_{P \text{ in } \mathbb{P}_i} \prod_{e \text{ in } P} a(e) \right] \left[\sum_{P \text{ in } \mathbb{P}_j} \prod_{e \text{ in } P} a(e) \right] \tag{10}$$

Equation 10 follows from the fact that all efficient paths can be constructed by independently choosing a member from each of $\mathbb{P}_i$ and $\mathbb{P}_j$ and putting link (i,j) in between and that all such combinations constitute efficient paths. Substituting Eq. 10 into Eq. 8, we see that

$$\text{prob}[(i,j)|j] = \frac{k\, a(i,j) \left[\sum_{P \text{ in } \mathbb{P}_i} \prod_{e \text{ in } P} a(e) \right] \left[\sum_{P \text{ in } \mathbb{P}_j} \prod_{e \text{ in } P} a(e) \right]}{\sum_i \left\{ k\, a(i,j) \left[\sum_{P \text{ in } \mathbb{P}_i} \prod_{e \text{ in } P} a(e) \right] \left[\sum_{P \text{ in } \mathbb{P}_j} \prod_{e \text{ in } P} a(e) \right] \right\}} \tag{11}$$

$$\text{prob}\,[(i,j)|j] = \frac{k\,a(i,j)\left[\sum_{P \text{ in } \mathbb{P}_i} \prod_{e \text{ in } P} a(e)\right]\left[\sum_{P \text{ in } \mathbb{P}_j} \prod_{e \text{ in } P} a(e)\right]}{k \sum_{P \text{ in } \mathbb{P}_j} \prod_{e \text{ in } P} a(e) \left\{\sum_i \left[a(i,j) \sum_{P \text{ in } \mathbb{P}_i} \prod_{e \text{ in } P} a(e)\right]\right\}} \tag{12}$$

$$\text{prob}\,[(i,j)|j] = \frac{a(i,j) \sum_{P \text{ in } \mathbb{P}_i} \prod_{e \text{ in } P} a(e)}{\sum_i \left[a(i,j) \sum_{P \text{ in } \mathbb{P}_i} \prod_{e \text{ in } P} a(e)\right]} \tag{13}$$

Arguing by induction that the forward pass of algorithm 1 calculates link weights, we can show that w(e) in the algorithm is equal to the numerator of Eq. 13 when e = (i, j). Hence,

$$\text{prob}\,[e|j] = w(e) \Big/ \sum_{e' \text{ in } F_j} w(e') \tag{14}$$

is just a rewriting of Eq. 13. The right side of Eq. 14 is precisely the quantity the algorithm uses to divert trips from node j to link e. Using Eq. 14 and again arguing inductively, but this time in the order to the backward pass of algorithm 1, we can show that, at the time that trips of node j are diverted, the quantity

$$y(j) = \sum_{e \text{ in } I_j} x(e) \tag{15}$$

does comprise all trips from o to d that are expected to go through node j. This completes the proof that the diversion volumes are indeed those implied by the path probability defined in Eq. 2.

The termination of the algorithm is obvious. Each link is processed twice, at most, and there are a finite number of links in the network.

The Parameter θ

To affect diversion probabilities and thus satisfy specification 5, the user sets the value of the parameter θ appearing in the exponent of the link likelihood a(e). As shown earlier, as θ varies from zero to infinity, the probability of using a particular path, which is Δt longer than the shortest path, is directly proportional to exp $(-\theta \cdot \Delta t)$. Thus, as θ increases, the probability that a trip will use a shortest path also increases. When θ is zero, all efficient paths are considered equally likely; the topological significance of a link in an efficient path is its sole criterion for attracting trips. At the other extreme, when θ is large, i.e., 10 or larger, the effect is a multiple shortest-path assignment, which assigns trips to all and only shortest paths. This allows the network designer to perform the equivalent of an all-or-nothing assignment that appropriately considers parallel routes.

Between these 2 (useful) extremes, presumably, there is a value for θ that does the best job in duplicating the results of human behavior. This writer does not know what this value is, or even whether a unique value would suffice for all o-d pairs, trip purposes, or geographic locations. This is a good subject for future experimentation. Given route selection data, we could estimate θ directly by using numerical curve fitting techniques. Alternatively, screenline interviews, in which trip-makers crossing particular links are asked the origin, destination and purpose of their trips, could provide a target toward which an iterative calibration procedure would aim. Or finally, θ would be tinkered with as the network analyst now tinkers with link times until the assigned volumes satisfactorily duplicated the observed ground counts.

Figure 5 shows inferred link volumes by using various values of θ in the preceding example, where we assigned 700 trips between nodes 1 and 25. The 4 numbers posted on each link are, from top to bottom, the link volumes obtained with the parameter θ equal to 0, 1, 2, and 10. Notice how, as θ increases, trips are drawn to the links on the shortest path.

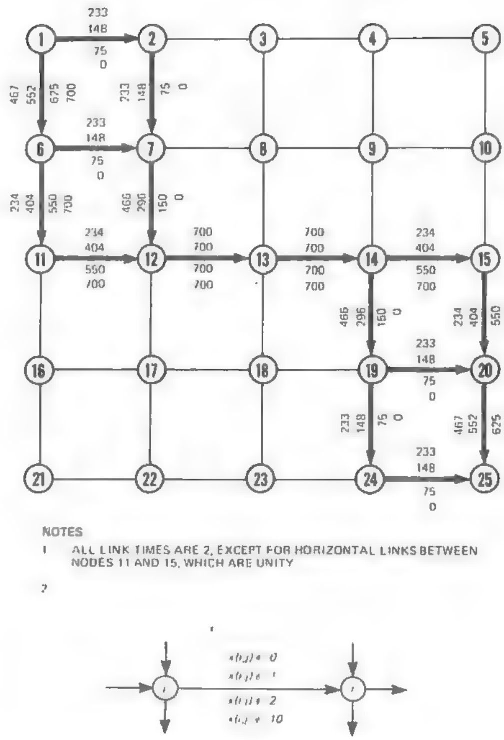


Figure 5. Arc volumes using various values of parameter θ

MODEL 2—PARALLEL PROBABILISTIC ASSIGNMENT

This second model redefines an efficient path to exclude constraint 2 above. In this alternative model, a path is efficient if all of its links satisfy constraint 1. This eliminates the need to know the shortest path distance from each node to the destination node d, and all trips originating at the origin node o to all destinations can be assigned simultaneously, in a single execution of algorithm 2. Thus, this second algorithm, a minor variation of the first, is effectively orders of magnitude more efficient than the first, but it is less discriminating in its selection of probable paths.

Efficient Path Redefined

In this alternative model, we discard one of the conditions for path efficiency and let a path be efficient if and only if all of its links have an initial node closer to the origin node than is its final node. Recall that this property was a necessary condition in the original definition of an efficient path. We now let it constitute a sufficient condition. Thus, all of the efficient paths under the original definition are contained in the set of efficient paths as presently redefined. We now have a larger number of efficient paths between a given origin and destination node pair than before. Thus, trips will generally be spread over more links than before.

Although the new definition of an efficient path yields more efficient paths, it surprisingly provides immense computational benefit. We assign to more paths, but the new definition permits us to do it a great deal faster. For example, in a network with 1,000 origins and 1,000 destinations, the number of executions of algorithm 2 would be 1,000 times fewer than the number of algorithm 1 with the old definition. Algorithm 2 would execute 1,000 times; algorithm 1 would execute 1,000,000 times. With the new definition, all trips from a given origin zone are assigned in roughly the same amount of time as was previously required for a single, widely separated origin and destination node pair. This fact becomes apparent in the following algorithm description.

The algorithm for the parallel assignment model is obtained by using the new definition of an efficient path and slightly modifying the multipath algorithm described earlier. The principal difference is that this parallel multipath assignment algorithm maintains node volumes. It is described in 3 major steps.

Algorithm 2

Preliminaries—To simultaneously assign all trips from origin node o to all destination nodes requires that the following 4 items be known for each node: y(i) = the number

of trips from node o terminating at node i; $p(i)$ = the shortest path distance from node o to node i; I_i = the set of all links starting at node i; and F_i = the set of all links ending at node i. Again, letting link $e = (i,j)$ have length $t(i,j)$, we can calculate the link likelihood of each link e.

$$a(e) = \begin{cases} \exp \theta \, [p(j) - p(i) - t(i,j)] & \text{if } p(i) < p(j) \\ 0 & \text{if otherwise} \end{cases}$$

(Notice that the old constraint $q(j) < q(i)$ has been dropped from the definition of $a(e)$, thus we eliminate the need to know $q(j)$, the distance between j and d.)

Forward Pass—By examining all nodes i in ascending sequence with respect to $p(i)$, their distance from the origin, we can calculate for each link e in I_i its link weight

$$w(e) = \begin{cases} a(e) & \text{if } i = o \text{ (the origin node)} \\ a(e) \sum\limits_{e' \text{ in } F_i} w(e') & \text{if otherwise} \end{cases}$$

When the destination node most distant from the origin node o is reached, the next step is undertaken (notice that this step is identical to the first step in algorithm 1).

Backward Pass—Starting with the most distant destination node, we can examine all nodes j in descending sequence with respect to $p(j)$. For each link e in F_j, the following 2 steps are undertaken:

1. A trip volume $x(e)$ is assigned to each link e:

$$x(e) = y(j)\, w(e) \Big/ \sum_{e' \text{ in } F_j} w(e')$$

2. The node volume at e's initial node i is increased by e's link volume:

$$y(i) \leftarrow y(i) + x(e)$$

When the origin node o is reached, the assignment is complete. All trips originating at node o have been assigned. (At this time, the node volume of the origin, $y(o)$, should equal the total trips originating at o. This constitutes a good error check in a computer implementation.)

Example—Figure 6 shows the results of a single execution of the multiterminal, multipath algorithm. The origin node is 1. The number of trips from node 1 destined for each node appears posted above the destination node. These numbers constitute one row of a travel demand matrix or "trip table." For example, the number of trips from node 1 to node 3 is 40; from node 1 to node 7, there are 0 trips; and at node 1 there are 20 intranodal trips, which will not be assigned to any links, but are included to show that they are judiciously ignored.

Posted below each line e is $w(e)$, its link weight as calculated in the forward pass. Above each link e is $x(e)$, its assigned link volume as calculated in the backward pass. Figure 6 shows the way all trips are spread from a given origin node through the network. The reader may attempt to duplicate these results by playing algorithm 2. Or he may prefer to refer to Dial's paper (70) that discusses and exemplifies the algorithm in much finer detail.

In a single execution of the algorithm, all trips originating at node 1 have been assigned to a total of 151 different paths. Of these 151 paths, only 57 end at nodes whose terminal volume is nonzero. Thus, 94 paths were "loaded" with zero trips. Of the 151 different paths originating at node 1, exactly 3 paths terminate at node 13. We may, for example, approximate the effective path assignments for the 40 trips from node 1 to node 13, as follows:

1. 23 trips were assigned to the shortest, 6-min path through nodes 1, 6, 11, 12, and 13;
2. 8.5 trips were assigned to the 7-min path through nodes 1, 6, 7, 12, and 13; and
3. 8.5 trips were assigned to the 7-min path through nodes 1, 2, 7, 12, and 13.

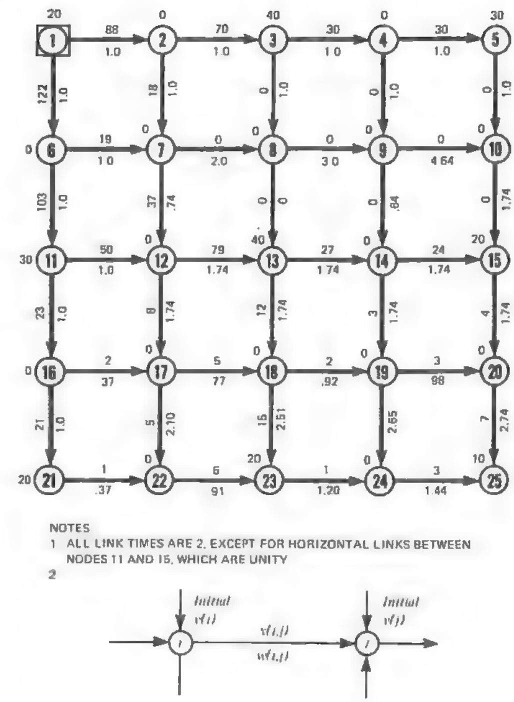


Figure 6. Parallel multipath assignment.

Table 1 gives the number of efficient paths to each destination node. General path-counting algorithms are given elsewhere; however, in this case, the preceding path volumes can be readily verified by referring to Figure 6, where the link weights are posted. Notice that all efficient paths from 1 to 13 go through node 12. Furthermore, only one of these paths, the shortest path, uses link (11, 12). Therefore, the probability of using the shortest path is identical to the conditional probability of using link (11, 12), given that node 12 is used:

$$\begin{aligned} \text{prob [shortest path]} &= \text{prob } [(11, 12) | 12] \\ &= w(11, 12) \Big/ \sum_{e \text{ in } F_{12}} w(e) \\ &= w(11, 12)/[w(11, 12) + w(7, 12)] \\ &= 1/1.74 \end{aligned}$$

Thus, the expected number of trips from 1 to 13 that will use the shortest path is $40/1.74 = 23$. The remaining 17 trips must split equally among the 2 remaining equal length, and therefore equiprobable, efficient paths. In the general case, of course, this calculation is not so simple, and reference should be made to Eq. 5 to determine individual path probabilities.

Justification of Algorithm 2

The effective difference between the faster, parallel assignment algorithm and its predecessor discussed in the preceding is in the parallel procedure's larger set of reasonable paths. For example, Figure 6 shows that, of the 10 trips from node 1 to node 25, 1 trip used paths that crossed link (23,24). Under the prior, more restrictive definition of a reasonable path, this link received no trips at all; there were no reasonable paths that used it. When we drop the explicit constraint that all nodes in a path must progress closer to the destination, we allow greater divergence of a path as it approaches its destination. Perhaps the tremendous computational advantage of the parallel multipath algorithm could be outweighed by an undesirable admission of too

TABLE 1

TRAVEL DEMAND AND EFFICIENT PATHS ORIGINATING AT NODE 1 AND TERMINATING AT NODE i

Destination Node i	Travel Demand[a]	Efficient Paths
1	20	0
2	0	1
3	40	1
4	0	1
5	30	1
6	0	1
7	0	2
8	0	3
9	0	7
10	0	11
11	30	1
12	0	3
13	40	3
14	0	3
15	20	3
16	0	1
17	0	4
18	0	7
19	0	10
20	0	13
21	20	1
22	0	5
23	20	12
24	0	22
25	10	35
Total	230	151

[a]Intranodal trips.

many paths. On the other hand, the parameter θ in the calculation of link likelihood may be used to effectively restrict path diversion. Judicious use of this parameter could render the multiterminal model useful and allow the economic benefit of its tremendous computational efficiency.

The utility of this alternative model rests on the practicality of the revised definition of an efficient path and the user's valuation of computer time. Otherwise, its properties are identical to the probabilistic assignment described previously. The new definitions do not invalidate the statements proved previously. Inefficient paths still receive no trips. Among efficient paths, the shorter ones still get more trips than the longer ones. All the highly desirable properties discussed earlier still hold; only the meaning of efficient path has changed. Formal proof of these statements is contained in the justification of algorithm 1. The only needed addition to the proof is an inductive argument showing that the node volume y(i) is properly maintained and is complete before being distributed to the links in F_i.

CONCLUSION AND RECOMMENDATION

This paper has briefly described and justified the mechanics of 2 efficient multipath traffic assignment models. It is shown elsewhere that the algorithms have extended utility. They can be readily modified to perform all-or-nothing assignment or

all-shortest-path assignment. They also can yield path statistics, revealing, for example, the length distribution of all competing efficient paths.

The models also appear promising in areas of "capacity-restraint" assignment and modal split. In capacity restraint, the model is an excellent tool to perform an incremental assignment, where in each iteration a fraction of the total trips are assigned and the speeds of the links are decreased to reflect their increased volumes. In modal split, it is intriguing to imagine using the model to divert trips among modes. In such a model, the network would be multimodal, and the diversion rules would appropriately reflect fare and transfer "costs" as well as preclude illegal mode changes en route. These 2 extensions are both interesting topics for further research.

As yet, the proposed model is an untested hypothesis. While laboratory experimentation using artificial networks has been encouraging, the model has not been tested by using full-scale, real-life transportation planning input. Therefore, it is recommended that such a test be undertaken to ascertain the model's utility.

While we prefer the aesthetics of the symmetric envelope of efficient paths of algorithm 1, we feel that most computers are not yet fast enough for its blanket substitution for all-or-nothing assignment in large networks. (It would, however, be quite feasible to restrict its employment to selected key origins and destinations, where it is felt that its path discrimination would be significant.) On the other hand, algorithm 2 is more efficient than some all-or-nothing model computer implementations. It would be practical to use it on a full-scale transportation network of 4,000 to 8,000 nodes. To this end, Alan M. Voorhees and Associates, Inc., under the sponsorship of the U.S. Department of Transportation, has completed a computer code for the parallel probabilistic assignment algorithm and the capacity restraint alluded to earlier.

This computer program is completely compatible with the Federal Highway Administration transportation planning programs for the IBM 360 computer. Its output include all the detailed information, e.g., turn volumes, that the planner has come to expect from such a program. With the program's input and output compatibility, the model can be put to work in a manner imposing no compromises on its user. He will have all the analytical and comparative capabilities present in the rest of the programs and, thus, possess an ideal mechanism to observe the model's performance in the field. The actual application of the model in the planner's workaday environment is, therefore, a feasible subject for further experimentation.

ACKNOWLEDGMENTS

There is a long list of individuals who have helped the writer in the development and description of these algorithms. At the top of the list are Walter Hansen, Howard Bevis, and Edgar Horwood. Support has also come from the U.S. Department of Transportation, the University of Washington, and Alan M. Voorhees and Associates, Inc.

REFERENCE

1. Dial, R. B. Probabilistic Assignment: A Multipath Traffic Assignment Model Which Obviates Path Enumeration. Univ. of Washington, Seattle, PhD dissertation 1970.

AUTOMOBILE USE PATTERNS IN NEW YORK CITY AND ITS ENVIRONS

J. David Jordan, Tri-State Transportation Commission

The transit revolution notwithstanding, the automobile will probably be the dominant means of personal transportation in the coming decades. Rates of daily automobile use in the Tri-State region surrounding New York City were related to several projectable variables in an effort to determine whether use would change significantly in the future. Rates of daily automobile trips, mileage, and time per automobile, which were obtained from home-interview survey data, were cross classified according to income, automobiles per household, persons of 16 years or older per household, and density. It was found that automobile use increased with automobiles per household in the low- and middle-income groups but decreased with automobile ownership within the high-income group. It was also apparent that automobile use increased with income and household size and decreased with increasing density. It was concluded, however, that because of the magnitude of these variations in use it is not necessary to incorporate them into the existing trip-forecasting model. A comparison of rates showed that, where use increased with automobiles per household in New York City, the opposite trend held in the surrounding suburbs. Daily automobile utilization in the suburbs was generally twice that in the city; it is recommended that trip forecasts for the 2 subregions be made separately.

•THE TRI-STATE region includes the 5 boroughs of New York City and 23 surrounding counties or planning areas. There are more trips made by automobile in this region than by any other mode of travel. This report first describes the characteristics of automobile trips made by the residents of the 3,600 square miles of the intensely developed portion of the region (Fig. 1) and then attempts to explain the variations in automobile trips, mileage, and time per automobile through relationships to income, population density, number of automobiles owned, and persons of driving age per household.

Data for this study were drawn from a 1963 home-interview survey, which consisted of a 1 percent sample of households in the intensely developed or cordon area. The results of the survey were validated at the county level with secondary source material.

The following numbers provide an overview of automobile ownership and use in the surveyed area.

VEHICLES IN TRI-STATE REGION

In 1963, there were more than 4 million private automobiles in the Tri-State study area where 16.2 million persons resided. Of the 5.3 million households in the area, 39 percent had no automobile, whereas 16.2 percent had 2 automobiles or more available:

Sponsored by Committee on Transportation Forecasting.

Figure 1. Tri-State cordon area.

Automobiles per Household	Households		Automobiles Owned	
	Number	Percent	Number	Percent
0	2,060,000	39.0	0	0
1	2,366,000	44.8	2,366,000	56.5
2	760,000	14.4	1,520,000	36.3
3+	96,000	1.8	301,000	7.2
Total	5,282,000	100.0	4,187,000	100.0

There was an average of about 4 persons per automobile inside the cordon, but this ratio ranged from more than 11 persons per automobile in Manhattan to between 2.5 and 3 persons per automobile in the suburbs.

TRIPS IN TRI-STATE REGION

Of the 29.6 million daily trips made within the intensely developed area of the region each day, more than 64 percent were made by automobile:

Mode	Unlinked Trips	
	Number	Percent
Automobile driver	13,374,000	45.3
Automobile passenger	5,641,000	19.0
Subtotal	19,015,000	64.3
Other modes	10,565,000	35.7
Total	29,580,000	100.0

There were, on the average, 3.19 automobile driver trips for every available automobile per day, and each trip averaged 4.4 airline-miles in length. In addition, for every automobile available, automobile driver-miles averaged 14.05, and automobile driver-minutes averaged 59.02.

These averages are characteristic when all the households in the cordon area are considered. However, in the subsequent sections on cross-classification analysis, only households with automobiles are considered and their corresponding averages are slightly lower. This was because about 120,000 automobile driver trips (0.9 percent of total) originated in households that did not have a private automobile available and could not be included because all the rates under study involved automobiles in the denominator. These trips were probably made with borrowed or rented vehicles.

OUTLINE OF ANALYTIC APPROACH

Three rates expressing characteristics of automobile use were examined in some detail.

1. Automobile driver trips per automobile—For any given area or grouping, the weighted average was derived as follows:

$$U = \frac{\Sigma \text{ automobile driver trips per day}}{\Sigma \text{ automobiles owned}}$$

2. Automobile driver-miles per automobile—Automobile driver-miles for a given trip was the straight-line distance between the origin and destination. For a given area or class, the following formula was used:

$$L = \frac{\Sigma \text{ automobile driver airline-miles per day}}{\Sigma \text{ automobiles owned}}$$

3. Automobile driver-minutes per automobile—Automobile driver-minutes for a given trip were minutes spent in the vehicle (exclusive of time spent walking) as reported by the driver. For a given area or class, the following formula was used:

$$T = \frac{\Sigma \text{ net automobile driver minutes per day}}{\Sigma \text{ automobiles owned}}$$

Variations in each of these rates were studied in the light of variations in the following discrete variables:

1. Median household income, I;
2. Automobiles available per household, A/HH;
3. Persons aged 16 years or older per household, P16+/HH (this was the best approximation available for those in the household eligible to drive); and
4. Density in terms of persons for each residential square mile, RD.

The effect of density on the rates was examined only perfunctorily for cross-classified data because the correct density for each group was not available and county density was used as an approximation. The effects of high density and an extensive transit system on automobile use were derived, however, through the comparison of the relationships in New York City data with those outside the city.

Two methods of analysis were applied to the data: (a) regression and correlation, and (b) cross classification. The first was applied cursorily to gain some perspective

of the strength of the association among the variables. The second method was pursued in greater detail and was reported in the following sequence:

1. The effects of income, automobile ownership, and potential drivers per household on automobile driver trips per automobile;
2. The effects of the same 3 independent variables on automobile driver-miles per automobile;
3. The effects of the same variables on automobile driver-minutes per automobile; and
4. Automobile use in New York City, in which the same approach was generally followed as in the first 3 steps, with New York City data contrasted with data from the area remaining within the cordon.

REGRESSION ANALYSIS

The study area was divided into 158 districts and an analysis of correlation was performed on household data aggregated to the district level. The dependent variable, automobile driver trips per automobile, showed significant linear correlation with each of two of the 4 independent variables, automobiles per household and the logarithm of density. The 2 variables were highway intercorrelated (R = 0.96) however; and automobiles per household emerged as by far the more significant explainer of variance in automobile use. The additional explained variance on adding density as the second independent variable was only 0.1 percent. In fact, automobile ownership emerged as the dominant explanatory variable when all 4 independent variables were regressed against automobile use, as shown in the following where automobile driver trips per automobile, U, is the dependent variable:

Independent Variables	Simple Correlation Coefficient	Multiple Correlation Coefficient	Beta Coefficient
A/HH	(+)0.85	0.86	0.739
log RD	(-)0.82		0.141
	(+)0.57		0.098
P16+/HH	(+)0.56		0.078

In the multiple regression analysis, income in any combination with automobile ownership showed an implied decreasing rate of automobile use with increasing income. This rather illogical trend cast doubt on the usefulness of the equation and led to the consideration of methods of analysis more suited to the discrete nature of the income variable. The results of the regression analysis also led to the belief that some of the relationships were nonlinear and would, therefore, be better analyzed by a method without linearity restrictions.

The relationship between automobile use and automobile ownership was further explored through a stratification of districts into those with net population densities of 25,000 or more and those that were less dense.

Least square lines and their associated correlations were derived within each of the density groups. The high-density group showed a slightly stronger relationship between use and ownership than had the unstratified data, but the low-density group displayed an insignificant correlation. The results of this analysis of automobile driver trips per automobile available versus automobiles per household are as follows:

Group	Density	N	R	Mean U	Mean A/HH
Unstratified	All	158	0.85	3.17	1.03
Group I	RD $\geq$ 25,000	60	0.88	2.20	0.59
Group II	RD $<$ 25,000	98	0.19	3.76	1.30

According to these results, automobile use increased up to a district ownership level of about 1 automobile per household and then leveled off to a use rate of between 3.5 and 4.0 trips per automobile per day (Fig. 2.)

Stratification according to density had much the same effect on the relationship of each of the independent variables to the dependent variable. The high-density group showed correlation comparable to those for the unstratified data, whereas the low-density group revealed insignificant relationships. This indicated that automobile use in suburban areas varied within a small range regardless of variations in automobile ownership, density, income, or potential drivers per household. It was recognized, however, that aggregation of data to the district level tended to obliterate much of the variation in the quantities under consideration. For example, it was impossible to describe characteristics of households owning 2 and 3 automobiles, because no district had that high an average automobile ownership. The aggregation effect and linearity restrictions involved in the regression approach led to further exploration of the data through cross-classification analysis.

CROSS-CLASSIFICATION ANALYSIS

The household is where most trips originate. This method of analysis assumed that automobile-owning households with similar social and economic characteristics would use their vehicles in a similar manner. The method was used in a descriptive role only; no measures of statistical significance were derived. An analysis of variance approach would have added more processing and computation than was warranted for a study of this scale.

Household data were cross classified within county, in 3 dimensions, as follows:

Median Income ($)	Automobile Owned Per Household	Potential Drivers Aged 16 or Over Per Household
Low, 0 to 3,999	1	1
Medium, 4,000 to 9,999	2	2
High, 10,000 and more	3+	3+

Thus, for each of the 22 counties within the cordon area, there were 3 matrices of automobile use rates: one each for trips per automobile, miles per automobile, and

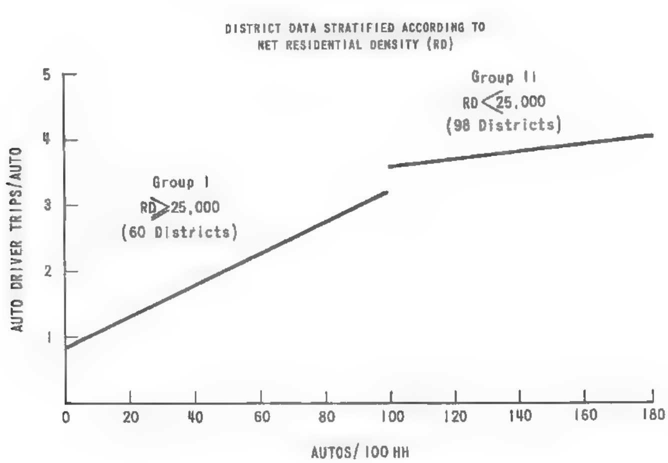


Figure 2. Automobile use related to automobile ownership.

minutes per automobile. Each matrix had 27 elements; for example, the first element would contain the weighted average automobile use rate for low income. households having 1 automobile, 1 resident aged 16 years old or older. The disadvantage at the county level was the insufficient sample sizes in many classes. For this reason, most of the analyses were conducted for the entire cordon area. The criteria for sufficiency were a minimum of 10 interviews. Samples were then inadequate in unusual classes, for example, where automobiles exceeded potential drivers in a household or where there were 3 or more automobiles per household in the low-income group.

AUTOMOBILE DRIVER TRIPS PER AUTOMOBILE

Automobile usage was defined by the number of automobile driver trips made by the residents of an area for each private automobile at the household:

$$U = \frac{\Sigma \text{ automobile driver trips per day}}{\Sigma \text{ private automobiles available}}$$

where U was 3.2 trips per automobile for the cordon area.

Effect of Automobile Ownership

Automobile ownership was expressed in 3 categories: households owning 1, 2, or 3 cars or more. When separated according to these categories, automobile use revealed a nonlinear relationship with automobile ownership (Fig. 3a). The average number of automobile trips per automobile was 3.17. Thus, the addition of 1 automobile to a household owning one automobile produced an increase in utilization of about 9 percent. The acquisition of a third automobile produced a decrease in use of about the same magnitude. The decrease in U after the second automobile is contrary to the results of the regression analysis described earlier, in which each additional automobile increased use by 2 trips a day. The difference in results can probably be attributed

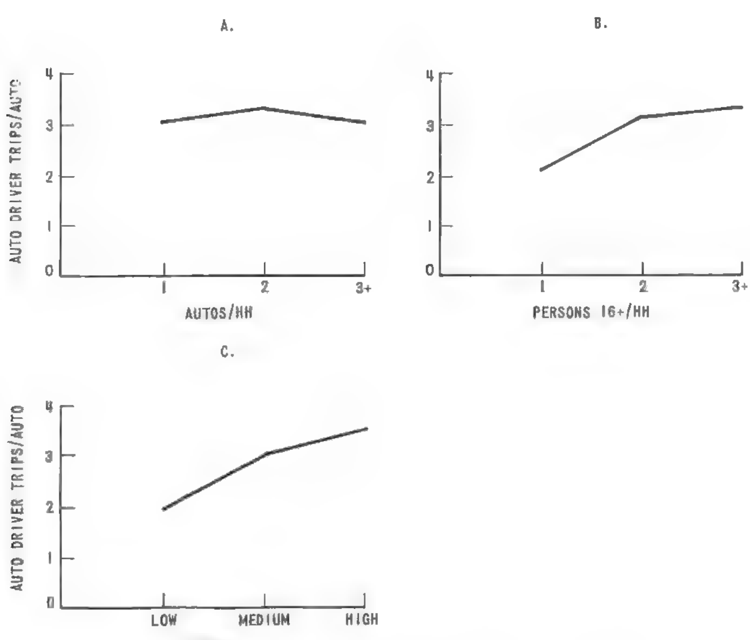


Figure 3. Automobile trips per automobile related to ownership, household size, and income.

to the bias introduced through the use of data compiled to the district level for input to the regression analysis. The highest average automobile ownership within the district was less than 1.75 per household.

Effect of Household Size

Household size was expressed as the number of people aged 16 years old or older in every household. This number should nearly equal the eligible drivers. The addition of 1 potential driver to the household with one driver increased automobile utilization by more than 1 trip. The addition of a third driving member to the household added about a fifth of a trip per automobile (Fig. 3b).

Effect of Income

When automobile use was derived within each of the 3 income classifications, it was found to increase with income. From low to medium income, there was an increase of more than 1 trip per automobile per day; from medium to high income there was a further increase of about half a trip a day per automobile (Fig. 3c). The relationship of automobile use with each of the 3 variables considered was nonlinear; use increased at a faster rate with the first increment than with the second for each of the variables. Although study of the separate effects of each of the variables on use had some value, all variables were found to be acting on use simultaneously, and it was this combined effect that was examined in detail.

Effect of Automobile Ownership, Household Size, and Income

A 3-way classification according to income, automobile ownership, and potential drivers per household produced a 27-element matrix of U-values. Five were eliminated because they contained fewer than 10 interviews, and all represented households where automobiles outnumbered potential drivers. The partial effect on automobile use of each of the 3 stratified variables was examined by changing 1 variable for all possible combinations of the remaining two.

Partial Effect of Household Size

If the income classification is disregarded, the effect of changes in the number of potential drivers on U was quite similar for households with 1 or 2 automobiles (Fig. 4d). Both ownership groups showed an increase of about 1 trip per automobile with the addition of a second driver and an increase of 0.15 to 0.20 trips with the addition of a third driver.

The pattern of variation in U with growing household size was distinct for the low-income group. Where the high- and middle-income use rates leveled off after the second potential driver, that of the low-income group continued to increase (Fig. 4a). This might indicate that people in these more wealthy households became passengers instead of drivers after the second adult; having already approached such a "saturation" level of trips as drivers (3 to 4 trips per day), they simply had little time left for driving. In the case of the low-income households, where U was low (under 2 trips per day), there was still time for the third driver to make a trip. The expense of owning an automobile would probably not be undertaken by a low-income family were the desire for an automobile not acute, and it would probably be used to the limit of the resources available for travel.

Partial Effect of Automobile Ownership

The effect on U of varying automobile ownership was quite small overall, but increased significantly within income class. The trend in automobile use was distinct for each of the groups; use increased with increase in number of automobiles in the low-income group and decreased with increasing number of automobiles in the high-income group (Fig. 5d). Automobile use rates were most dissimilar by income group for households owning only 1 automobile, and automobile use by the high-income group

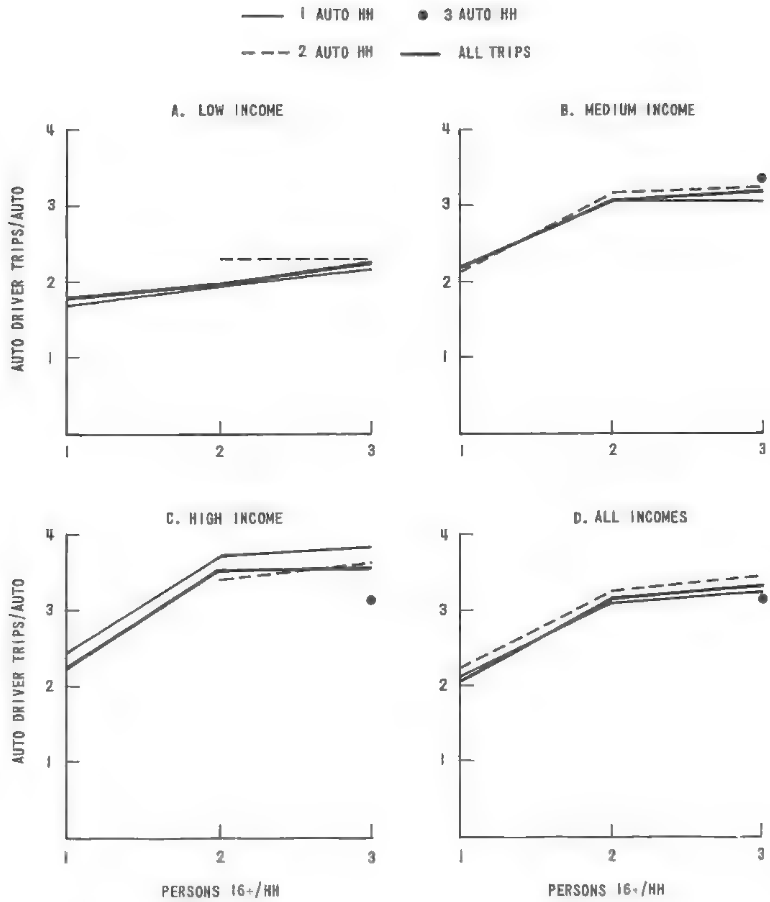


Figure 4. Automobile trips per automobile and potential drivers per household.

was almost twice that of automobile use by the low-income group. For households that owned 3 automobiles or more, the use rates were within 5 percent of one another for all income groups.

Once again, the results seemed to imply that as long as U was less than about 3 to 4 trips per day, the second and third automobiles would be used as much or more than the first. When that rate was reached, the additional automobile was used at the same or lesser rate than the first. The only exceptions to this were the supersaturated groups that had more automobiles than potential drivers.

Effect of Density

With stratified household data, it was not possible to obtain population densities because there was no way to assemble land area for each socioeconomic level. A rough approximation of the effect of density was attempted by assuming that the average county net population density would prevail regardless of the economic level of the segment under consideration.

Trip rates per automobile were related to net density according to income level, controlling automobile ownership, and household size. Taking the most populous sector—2-person households with 1 automobile—a fair correlation was obtained between U and the logarithm of density with each income stratum (Fig. 6). The higher the income level was, the faster U decreased with increasing density, although the percentage rate

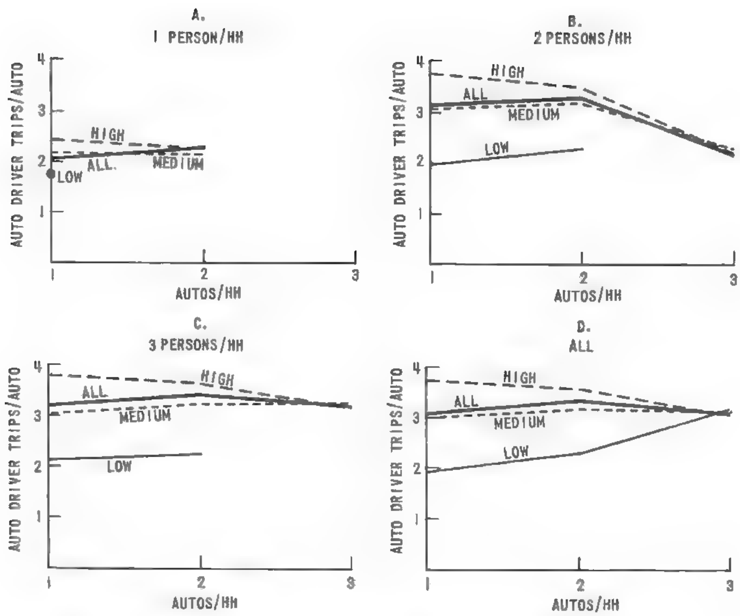


Figure 5. Automobile trips per automobile and automobile ownership.

of decrease in U was about the same in each income group. U varied with income most in the outlying counties and progressively less toward the central area. For automobile trips per automobile by households having 1 automobile and 2 persons aged 16 and over, the results were as follows:

Income	Density		
	Low	Medium	High
Low	2.90	2.28	1.21
Medium	4.59	3.54	1.72
High	5.83	4.56	2.37

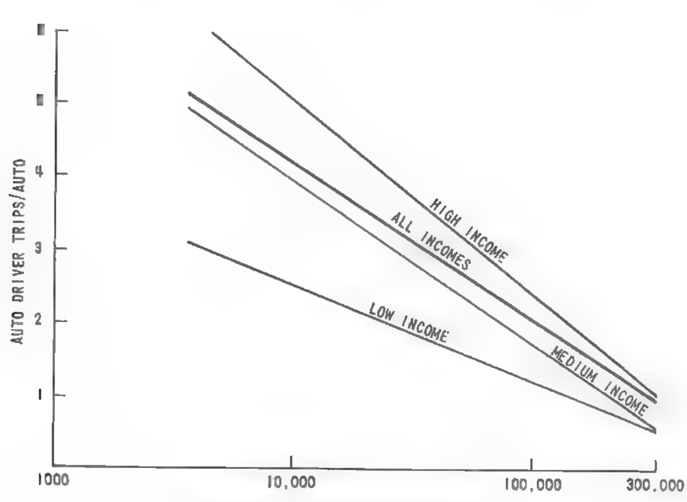


Figure 6. Automobile trips per automobile and density.

TRIP LENGTH—DISTANCE

The average airline-miles driven per automobile owned was computed for each group in the matrix:

$$L = \frac{\Sigma \text{ automobile driver-miles}}{\Sigma \text{ automobiles}}$$

where L is the mileage per automobile per weekday.

The average L for the cordon area was about 14 miles. Among the subcategories, L ranged from about 6.9 miles (1 person, 1 automobile, low income) to 16.8 miles per automobile (3 persons, 2 automobiles, high income). In general, changes in the independent variables caused automobile mileage to vary in much the same manner as the trip rate per automobile (compare Fig. 7 with Fig. 3); there was little variation in average trip length.

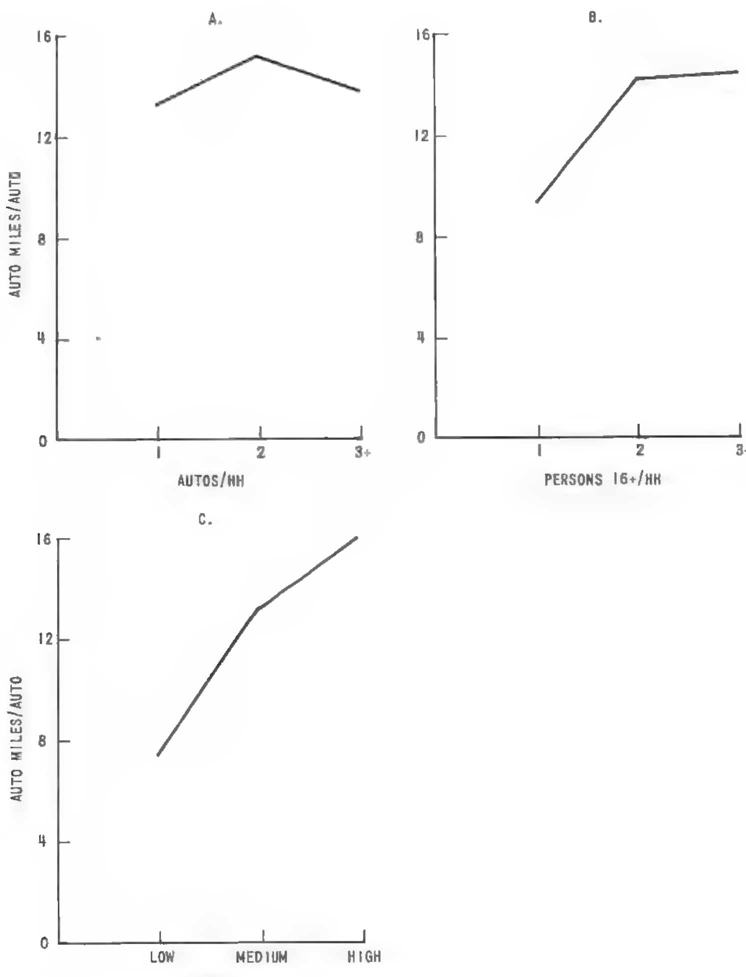


Figure 7. Automobile miles per automobile related to ownership, household size, and income.

Automobile Ownership

The second automobile acquired caused an increase in mileage for each automobile of about 14 percent (Fig. 7a). The third automobile brought the mileage down by about 9 percent. This meant an increase in trip length (airline-miles per trip) from 4.3 miles for households with 1 automobile to 4.5 miles for households with 2 and 3 automobiles, which was a difference of less than 5 percent.

The slightly unexpected phenomenon of increasing mileage per automobile with increasing automobile ownership was also found in Chicago, where households with 2 or more automobiles averaged 12.4 miles per automobile compared to 11.6 miles for households with 1 automobile (2). The increased use was most evident in the 3 densely populated central rings.

Another study conducted by the state of Illinois found that households with 1 automobile averaged 9,900 miles per automobile per year compared with 10,000 miles for households with more than 1 automobile (3). A national study of vehicle use showed that households with more than 1 automobile averaged 9,300 miles per automobile per year compared to 8,900 miles for households with 1 automobile (3).

Of the studies mentioned, only the Tri-State data displayed an increase in mileage per car of more than 10 percent in the transition by households from 1 to more than 1 automobile. The greatest increase was observed within Tri-State's low-income group and Chicago's inner rings, both of which had mileages well below their respective area-wide averages for households with 1 automobile and moved closer to the area-wide averages for households with 2 automobiles.

Work-Trip Length

There was some evidence to support the contention that the second (probably the newer) automobile was driven farther to work than the first. Seventeen of 22 counties in the cordon area had a longer automobile driver work-trip length for households with 2 automobiles than for households with 1 automobile. The cordon area average for households with 2 automobiles was 6.0 compared to 5.6 airline-miles for households with 1 automobile. Whether the 7 percent difference in trip length was statistically significant was not determined. Average trip lengths for trips bound for other destinations showed no such trend.

The transition from 2 to 3 automobiles per household brought a decrease in work-trip length in 15 of the 22 counties; however, the sample of households with 3 automobiles was small. The average work-trip length for households with 3 automobiles was 5.7 airline-miles. It would appear possible, then, that the purchase of the second automobile was often tied to the job location being farther away or the new home being farther from the job. Considering the increase in use per automobile discovered earlier, the latter was more likely because the move was probably to a more automobile-oriented community.

Household Size and Income

Similarly, the effects of household size and income on mileage closely paralleled their effects on trips per automobile (Fig. 7). The household with 2 persons 16 years old or older, more likely to be a family unit than a household with 1 person over 16 years old, had more needs and therefore traveled more to satisfy them. The addition of the third adult had little effect on the mileage per car. Household size had no significant effect on trip length (miles per trip).

The low-income group accumulated 7.5 miles per automobile, compared with 13.4 miles per automobile for middle-income residents and 16.1 miles per automobile for the high-income group (Fig. 7c). The latter could afford to take advantage of opportunities that were distant from home and thus accumulated more than twice the mileage per automobile of the low-income group. However, the difference in mileage per automobile driver trip between the 2 extreme income groups was less than 20 percent.

Effect of 3 Independent Variables

Having again cross classified according to automobile ownership, persons 16 years of age and more per household, and income level, we found the combined effect of these variables on mileage per automobile to be similar to their effect on the trip rate per automobile; that is, mileage per trip (trip length) did not vary significantly.

TRIP LENGTH—TIME

The average time each automobile was driven per day was computed for each social or economic group:

$$T = \frac{\Sigma \text{ net automobile driver-minutes}}{\Sigma \text{ automobiles}}$$

where T is the net minutes per automobile per weekday.

The average for the cordon area was 58.4 min, ranging from 35.2 min (2 persons, 2 cars, low income) to 71 min per car (3 persons, 1 car, high income). Trip time per

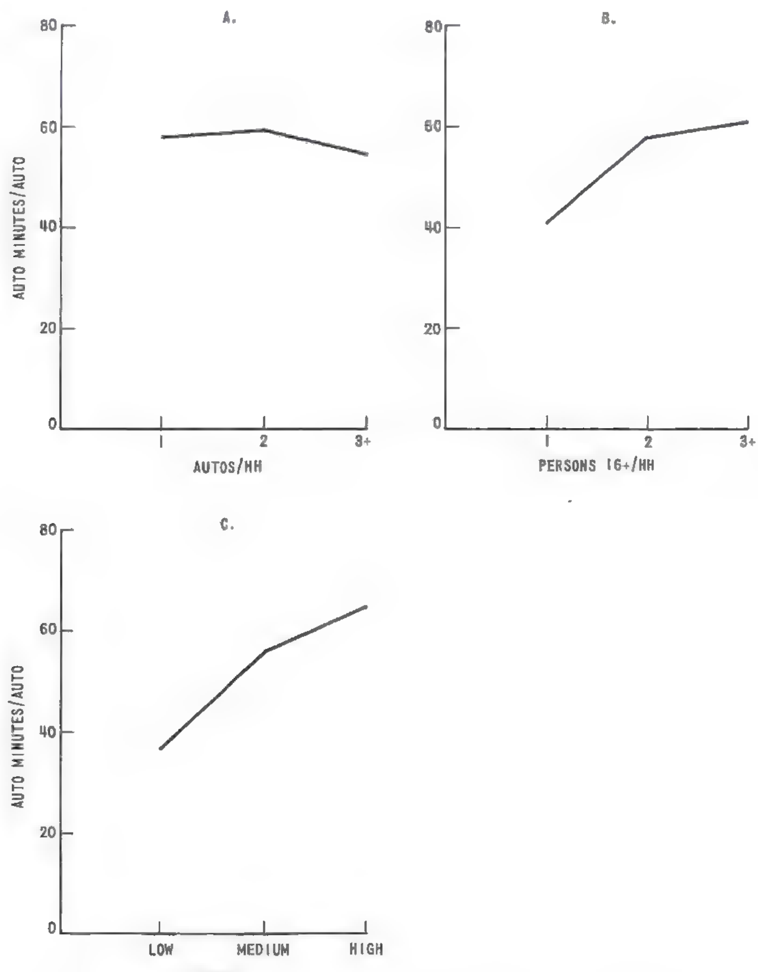


Figure 8. Automobile minutes per automobile related to ownership, household size, and income.

trip (18.4 min for the cordon area) did not vary much, and, in general, the magnitude and direction of changes in T with increments in the independent variables were similar to those for U and L (compare Fig. 8 with Fig. 3).

AUTO USE IN NEW YORK CITY

The average automobile use rate was 1.97 trips per automobile for New York City residents, 3.66 trips for residents of the remainder of the intensely developed area, and 3.17 trips for residents of the study area. This reflected, in part, the lower income level and smaller household size, in general, in the city. In addition, the extensive transit system and the close convenience of shopping and other facilities, along with the high cost of highway congestion and parking, appear in many cases to have kept the rate of automobile use low on weekdays. Thus, about a quarter of the trips by residents of automobile-owning households in the city were by subway and, altogether, 47 percent by modes other than the automobile, compared to only 16 percent in the suburbs.

First, when the independent variables are considered separately, the incremental effects of potential drivers per household and income on automobile use were similar for both New York City residents and nonresidents. Although the initial U was lower in New York City by at least 1 trip, comparable increases in potential drivers or income brought comparable increases in U for city and noncity (Figs. 9b and 9c).

This was not the case where automobile ownership was concerned. Households with 1 automobile in the suburbs generated more than 2 trips per automobile more than those in New York City. The addition of a second automobile caused the automobile use rate to fall to 3.5 trips per day in the suburbs and to rise to 2.3 in the city (Fig. 9a). The third car caused a further decrease to 3.1 trips in the suburbs and a slight decrease to

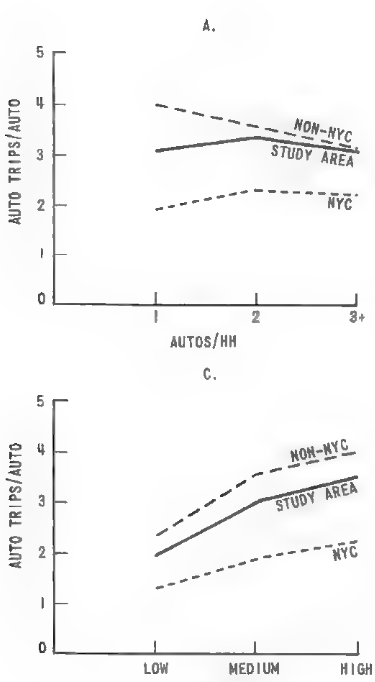


Figure 9. Automobile trips per automobile related to ownership, household size, and income—New York City comparison.

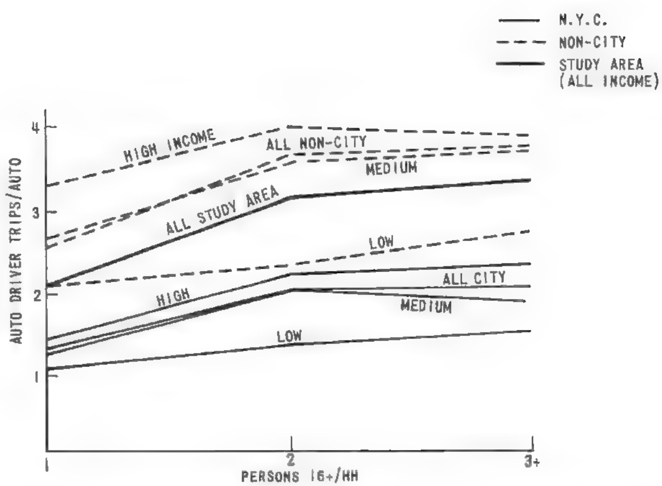


Figure 10. Automobile trips per automobile and automobiles per household stratified by income level—New York City comparison.

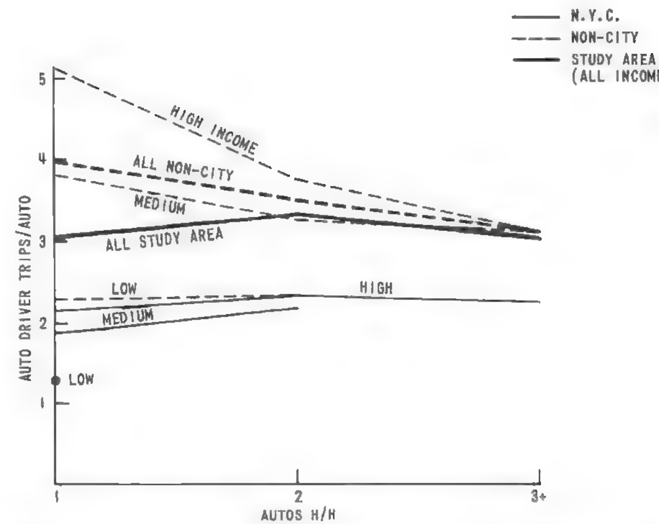


Figure 11. Automobile trips per automobile and persons 16 years old and more per household stratified by income level—New York City comparison.

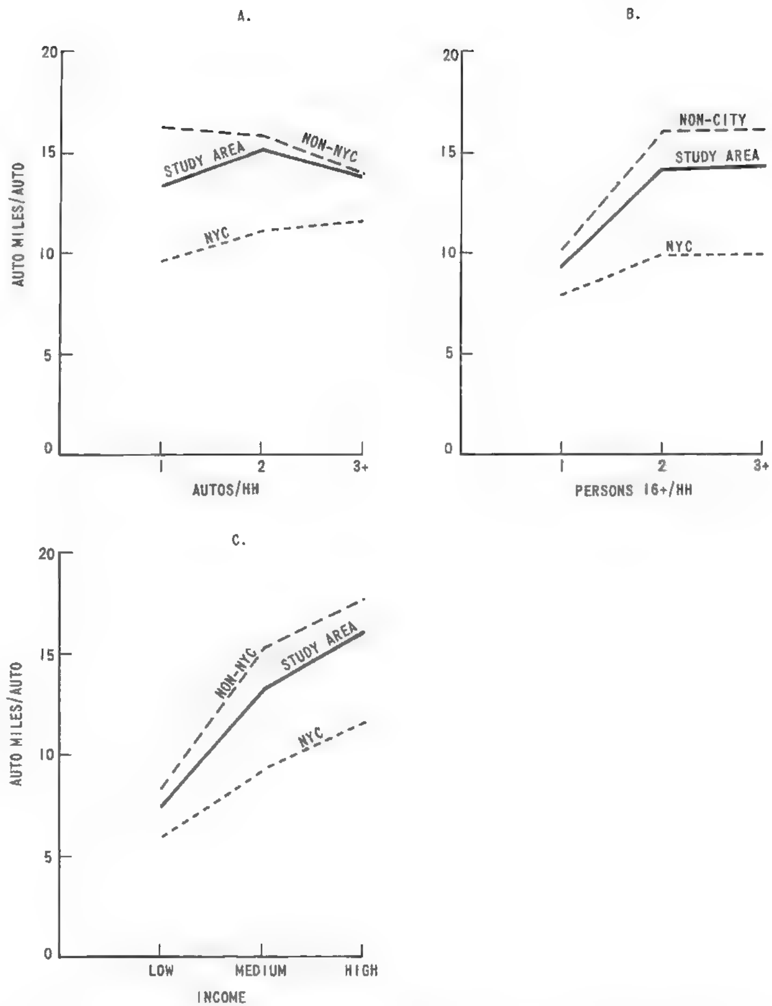


Figure 12. Automobile miles per automobile related to ownership, household size, and income—New York City comparison.

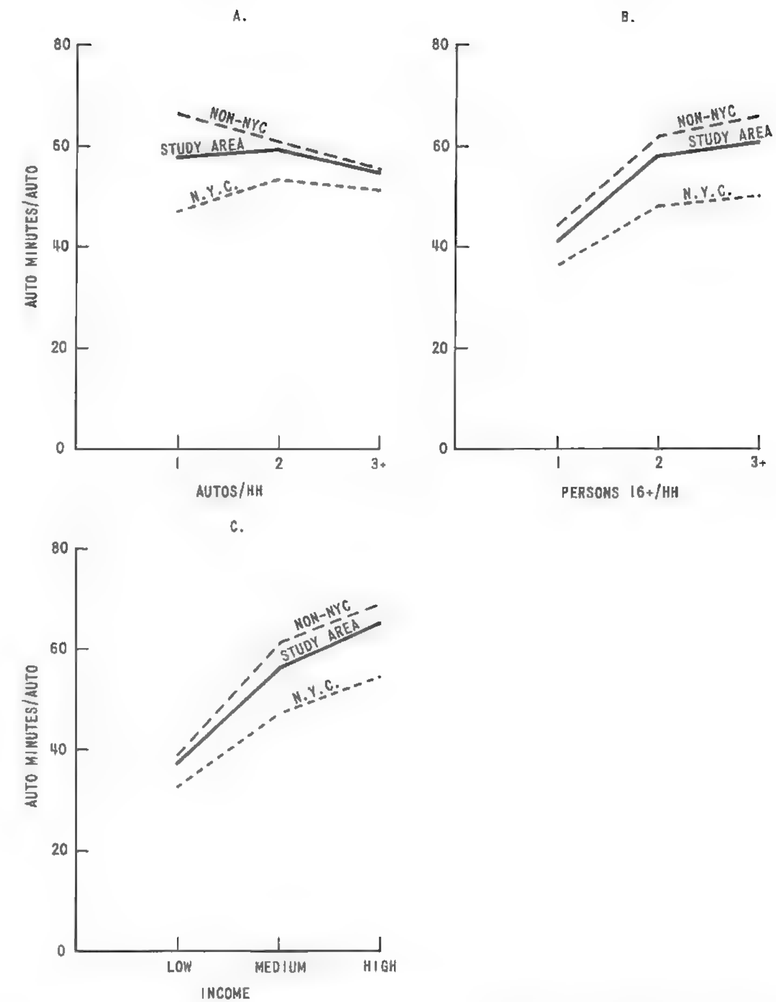


Figure 13. Automobile minutes per automobile related to ownership, household size, and income—New York City comparison.

2.2 trips per automobile in the city. (The number of households with 3 automobiles or more in the city was relatively small.) These opposite trends were similar to those found for cordon area high- and low-income groups described earlier; the use rates for households with both 1 automobile and 2 automobiles were almost identical in the city and in the low-income stratum of the cordon area.

Household Size and Income

The effect of increasing household size was similar in the city and remaining area. The suburbs generated 1 to 2 trips more per automobile than the city at each level, but the common trend was an increase in automobile utilization of about a trip with the addition of the second potential driver, and then a much smaller increase in use with the addition of the third person over 16 (Fig. 10). The exception to this trend in both city and suburbs was the low-income group, which exhibited no leveling off of use with the addition of the third driver.

Automobile Ownership and Income

In the suburbs, U decreased with increasing automobiles owned within the high- and middle-income groups to about 3 trips for each automobile per day. The noncity, low-income group, however, showed use rates similar to the city's high-income group and an increased rate of use with increasing automobile ownership, also coalescing to about 3 trips per automobile (Fig. 11). This low-income stratum could hardly be classified suburban as it was probably composed largely of households in cities such as Newark and Jersey City and was subject to much the same influence as households in New York City.

High- and medium-income households in the city were very close in use rates. The low-income group had a low rate of use but the sample was small.

Trip Length (Miles) in New York City

The automobile belonging to an average household outside of New York City was driven 15.9 miles a day; this was 61 percent more than was driven by a similar vehicle that was garaged within the city and driven for an average of 9.9 miles. The average automobile trip, however, was slightly longer within the city; 5.0 miles to 4.3 miles. This held within all the subclass variations and probably meant that city drivers used their cars mainly for getting to special objectives that could not be reached easily by walking or by using public transit. While automobile ownership for each household increased, mileage per automobile increased in the city and fell in the remainder of the cordon area, generally following the trip-rate trend for the 2 subregions (compare Fig. 12 with Fig. 9).

Trip Length (Time) in New York City

The average city automobile was driven 48 min a day compared to 63 min for the noncity car. Congestion contributed to a slower trip in the city; airline speeds were 12.3 mph there, and 15.2 mph outside. Thus, the time elapsed on each trip was 24.4 min in the city and 17.7 min beyond its boundaries. Because trip time did not vary appreciably, minutes per automobile behaved in much the same manner as trips when the independent variables were varied (compare Fig. 13 with Fig. 9).

RESULTS AND DISCUSSION

Because miles traveled and time elapsed for each trip varied so little, automobile use generally refers to automobile driver trips for each car, mileage per car, and net minutes for each car.

For the cordon area as a whole and the effect of each independent variable on automobile use taken in turn, the following observations can be made:

1. Automobile use was about 9 percent higher for households with 2 automobiles than for households with 1 and 3 automobiles, for which the rate was about 3.1 trips

per automobile per day. The increase in use with the second car was limited to the medium- and low-income groups for whom, it was conjectured, the decision to purchase the additional vehicle would be made only if there was a significant desire, and its intensive use was therefore ensured. High-income households were more likely to purchase pleasure cars whose use would not exceed that of the existing car.

2. Automobile use increased with the number of potential drivers in the household but tended to level off at about 3.3 trips per automobile after the second person 16 years of age or over. Greatest use of the vehicle was made by the first 2 adults, presumably the parents.

3. Automobile use increased with income, leveling off as income increased beyond the $10,000 per year group. The high-income group, at almost 3.6 trips a day, used the automobile 80 percent more than the low-income group.

4. Automobile use decreased with increasing density within each income group.

Automobile utilization in New York City compared to that in the less dense areas revealed the following:

1. Households with 2 automobiles in New York City generated more trips per automobile than households with 1 automobile. The opposite was true outside the city. The use rate outside the city, at almost 4 trips per automobile, was about twice that inside the city for households with 1 automobile.

2. Income and household size affected usage in much the same way both inside and outside New York City. Use rates per automobile outside the city were generally 1 to 1.5 trips higher for any class or group.

IMPLICATIONS FOR TRIP FORECASTING

Although there were differences in automobile use patterns for different segments of the population, their application to the derivation of automobile trip estimates for future years is of questionable value. The magnitude of the differences in usage rates is probably smaller than the range of error involved in the projection of population and income. Even if it were assumed that the use rate would remain steady over time in each socioeconomic class, the degree of accuracy is doubtful in projections of households grouped according to automobile ownership, income, and household size.

ACKNOWLEDGMENTS

Invaluable aid was received in the preparation of this paper from several persons: general guidance and direction from Lawrence V. Hammel and Edward F. Sullivan; technical assistance from William H. Coelln, Thea Edelstein, and Elizabeth Phifer; editorial and clerical assistance from Lyman Coddington, Jerry Kendall, and Diane Sydnor; and data processing by Jocelyn Bishop, George Fasciana and Widge Connolly.

The preparation of this report was financed in part through federal funds made available by the Federal Highway Administration and an urban planning grant from the U.S. Department of Housing and Urban Development, under the provisions of Section 701 of the Housing Act of 1954, as amended, and in cooperation with the states of Connecticut, New Jersey, and New York.

REFERENCES

1. Oi, W. Y., and Shuldiner, P. W. An Analysis of Urban Travel Demands. Northwestern University Press, 1962, 281 pp.
2. Chicago Area Transportation Study Final Report, Vol. 2, p. 76, 1960.
3. Bostick, T. A., and Greenhalgh, H. J. Relationship of Passenger-Car Age and Other Factors to Miles Driven. Highway Research Record 197, 1967, pp. 25-35.
4. Mills, F. C. Statistical Methods. Henry Holt and Company, New York, 1955.

CHOICE OF ROUTES ON URBAN NETWORKS FOR THE JOURNEY TO WORK

Manfred H. Ueberschaer, School of Civil Engineering, Purdue University; and Greater Lafayette Area Transportation and Development Study

About 13,000 drivers traveling from home to work were interviewed for their daily routes on the highway network and for their motivations for selecting these routes. The traffic conditions and characteristics of each route were also investigated. The results of these investigations and an analysis of the motivations showed that travel time, distance, number of possible stops, and maximum lane volume on a link between the points of choice are important for route selection. These 4 resistance factors were used to develop a trip-diversion model. It was found that the parameter values of this model depend on route distance or driving time and the type of town district traversed. Optimal routes in highway networks were found by determining a formula for link resistance. The formula contains travel time, length, and lane volume of the link as important factors, and it is shown that the parameters of this equation are a function of the link length. An application of the model showed good results.

•THE FACTORS AND REASONS that influence a driver to take a certain route in a road network, if he has a free choice among several routes, are numerous and are dependent on the driver's mentality, attitude, social status, age, sex, and other factors. Especially important, however, is the trip purpose. As in other problems of urban transportation planning, the easiest trip purpose to handle is the trip from home to work because its travel pattern is very homogeneous on all working days. Characteristics of this trip purpose are as follows:

1. The origins and destinations of commuter trips are well defined, as shown by all regression analyses and calibrations of trip-generation and trip-distribution models;
2. Commuters on their way to work from home drive in a very direct, straight, and purposive manner, and they take routes by which they reach their destinations promptly;
3. Commuters have a good knowledge of the area and know alternative routes; and
4. Traffic patterns and conditions (volumes, travel times, and waiting times) during morning peak hours are usually the same and have low variances compared with the variances of volumes and travel times during other hours of regular working days.

The route choice of trips for another purpose (e.g., shopping or social-recreational) is not so certain. These routes may not be so well defined because of the more leisurely nature of the trip or because the trip destination is often casual or unknown.

Many reasons for route choice for all trip purposes, however, can be classified into 2 groups—road characteristics and traffic conditions. Some of those belonging to the first group are route length, road width and number of lanes, pavement conditions, design, hills, sight distance, speed limits, right-of-way, traffic control, railroad crossings, road construction areas, and scenery. Some factors related to traffic conditions are travel time, waiting time, speed, volume, commercial traffic percentage, public transit

Sponsored by Committee on Transportation Forecasting.

in street, and pedestrian crossings. Not all of these factors are measurable by means of length, time, cost, and other quantities. Many of them can be evaluated only by psychometric methods and often describe primarily the convenience or strain of a route. These more subjective reasons of drivers can be obtained only from interviews.

Although the factors affecting travel patterns of the trips from home to work during the morning peak period are rather homogeneous, the drivers coming from the same residential area and going to the same industrial area do not all take the same route. The route 1 driver takes depends on the criteria of route resistance he uses for the evaluation of all possible routes for his trip to work. Every driver uses different criteria objectively or subjectively. Even if all drivers would choose the same criterion, they would not take the same route, because they cannot evaluate the route resistance of all possible alternates objectively. Another reason is that most of the route criteria are not constant or unchangeable; they alter during time and can only be estimated. The result is that in most cases there is more than 1 possible reasonable route between any 2 traffic zones.

The aim of this study was to determine the criteria drivers use in choosing for the routes for their trips to work, criteria that are practicable and applicable for a diversion assignment model. Other goals were to define a resistance criterion for locating optimal routes by computer analysis and to develop boundary conditions for a realistic evaluation of alternative routes.

DATA COLLECTION

The needed information was obtained by interviewing of 13,000 employees who drove every morning from home to work in their cars. The interviews were made in 1965 in the following 5 of the larger German cities:

City	Population	Residents per Vehicle
Aachen	178,000	6.1
Bochum	367,000	6.7
Düsseldorf	706,000	5.3
Frankfurt	696,000	4.2
Leverkusen	103,000	5.3

These values are for the area within the city limits. However, the total surrounding area of influence of these cities, relative to commuters, is much greater. For instance in Bochum, 4.1 million people can reach the core of the city within 30 min. In Düsseldorf and Frankfurt the figure is about 2.5 million people. The interviewed people were, in most cases, blue- and white-collar employees of manufacturing industries and, in some other cases, employees of the commercial or business industries.

The interview form used was distributed in most cases by the employer's personnel office. It had on its front page a short information section and instructions on how to complete the form. The questions were easy and quick to answer without the help of an interviewer. As expected, many people refused to answer. The return rates ranged from 25 to 62 percent; however, about 13,000 completed questionnaires were returned.

The respondents were required to answer the following questions:

1. Where do you live (address and city)?
2. Describe the route you take every morning from home to work by giving 3 or 4 names of streets, bridges, and buildings located along the route. (The route description on almost all returned forms were precise enough for the route to be drawn accurately on a map. This method of route description was chosen because pilot interviews had shown that many people could not describe their path on a map.)
3. Why do you take this route? From the following 9 given motivations, check the 2 (but only two) most important to you: (a) only 1 route exists; (b) I drive it out of habit;

(c) it is the fastest route; (d) it is the shortest route; (e) it has good road design and pleasant scenery; (f) it has greater safety because of traffic control; (g) it has right-of-way at most intersections; (h) it has less congestion for the longest part of the route; (i) it is less hilly; and (j) there are other reasons not stated here.

4. When do you start from home each morning? This question was asked to obtain the peak period of the investigated home to work trips.

5. What is the travel time from home to work you usually require?

6. Where do you park your car while at work? From this question, the exact destination in the industrial area could be located.

7. Do you take the same route back to your home (answer yes or no)? A detailed description of this route was not asked. However, a special investigation in Frankfurt showed that 85 percent of the drivers took the same route on their way home.

The origins and destinations and routes obtained from this interview were located and analyzed with the help of a map. The most important interchanges and their given routes were used for a very intensive survey of the existing road characteristics and traffic conditions. For each road section or each intersection of the considered route, the following factors describing the road characteristics were determined: length, effective roadwidth, number of lanes in the considered direction, streetcars without separate tracks, grade, traffic control at intersections, cycle and phases of signalized intersections, coordination of signalized intersections, speed limitation, and location within the urban area. Factors concerning the traffic conditions that were obtained were mean speed, travel time, waiting time, volume per lane and direction, and truck percentage. These factors were gathered by local measurements and traffic observations during the morning period in the direction of the considered traffic flow and were checked by direct travel time measurements from origin to destination by test trips. The analysis of the interviews later showed that the average travel time stated by the drivers was almost always higher by 5 to 10 min than the measured travel time average and was independent of the trip length. The probable reason is that drivers include their walking time to and from the parking lot.

ANALYSIS OF INTERVIEWS

Number of Chosen Routes

The 5 investigated cities have somewhat different characteristics. Düsseldorf and Frankfurt are typically metropolitan in their residential areas and have very active central business districts and large industrial districts that sometimes reach to the city core. Bochum has a similar land use; however, it has 2 big neighbors, Essen and Dortmund, where many of the interviewed drivers came from. It also has a very efficient highway and freeway system. Aachen has major industries in the northern and eastern fringe area, and, because many employees come from neighboring communities to the north or east by using rural highways or the autobahn, its traffic pattern is different from that of the other cities. Leverkusen is located at a very important autobahn intersection and has a good highway network to its neighborhoods, where most of its employees live.

These short descriptions of the investigated cities, their highway networks, and relationships to their surrounding neighborhoods are made to explain the following findings. It was found that the number of routes between origin and destination is dependent on the following:

1. The trip length measured by the airline distance. With increasing distance, the number of selected routes increases, having a maximum of 3 to 5 routes at a 2- to 6-mile distance. Then, with increasing distance, the number of alternatives decreases to 1 or 2 routes because the drivers for longer trips apparently prefer freeways or fast highways and stay on them to the exit nearest their destination.

2. The type of highway network the drivers mainly used. The number of accepted routes is lower in a rural highway network than in a typical urban network. Autobahns or other freeways are generally preferred and reduce the number of chosen routes in an area more than when no freeways are available.

3. The district of the town that is traversed. In central business districts, the number of alternative routes is higher than in fringe and suburban areas, depending on the arterial network density.

4. The traffic load of the arterial road network.

One may note close interrelationships among the last 3 items. However, the combinations of them in the 5 investigated cities were quite different and had an effect on the number of chosen alternative routes. For instance, Frankfurt has a higher number of alternative routes than Aachen does because the characteristics previously described are different (Fig. 1). Bochum's widely spread distribution is a good indication for its location in an intensely urbanized neighborhood.

Alternative routes were considered as only those that were used by more than 3 drivers in the sample during the morning peak hour. If one included the routes taken by only 1 driver, the alternatives would have been 2 to 3 times as high (white bars in Fig. 1). Such routes were regarded in further considerations as not indicating the behavior of a significant number of drivers, even if they were logical routes.

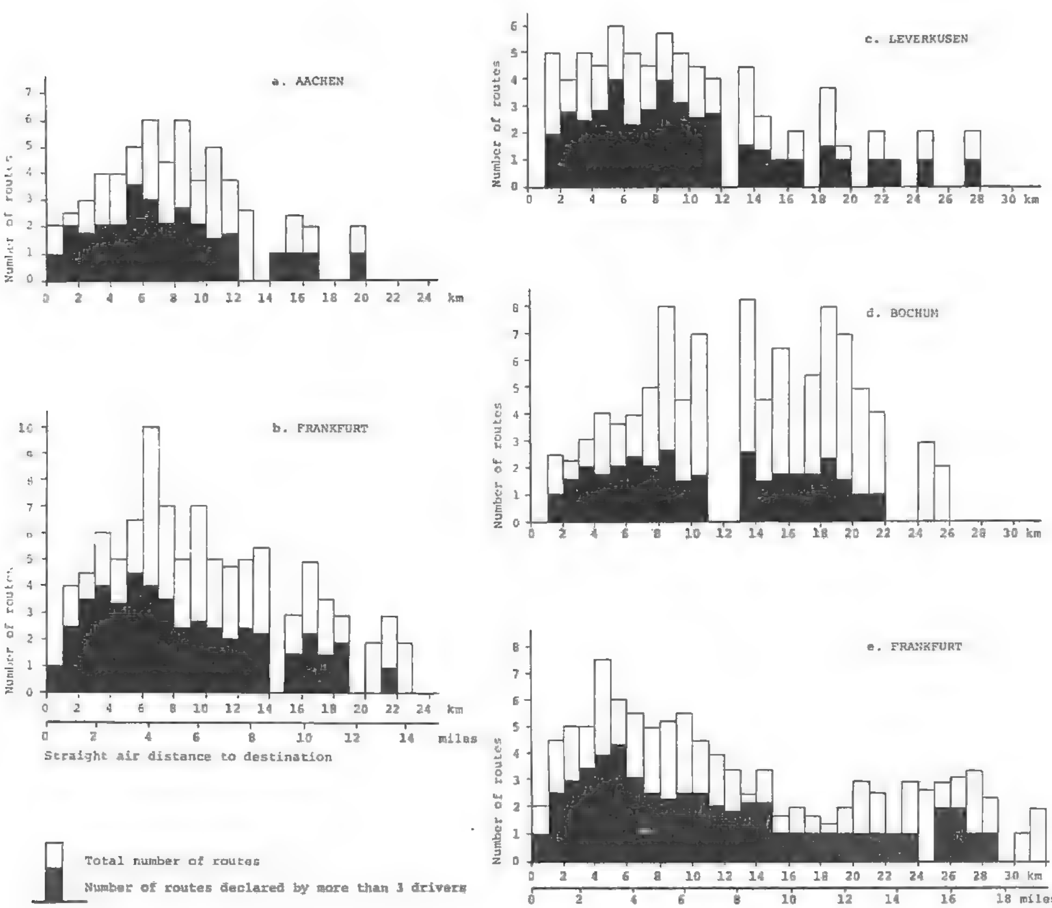


Figure 1. Alternative routes versus length of trip (air distance).

Motives for Route Choice

The questionnaire was simplified for the respondents in that it had 9 allowed statements about route-choice reasons. The one or two that seemed to be most important to the driver in his preference of route were to be checked. These statements were chosen from pilot studies conducted previously. Item 10 gave an option to the respondent to name another reason not listed in the other items, but it was not used very often. When it was used, only special reasons (no railroad crossings or better access to a freeway) were named, or else it was used for complaints about very bad traffic situations at special points in the network.

Most commuters in all 5 cities took their chosen route because they believed it to be the shortest in time; the second most important reason was that it was the shortest route in distance. The next important reasons, but not nearly as important as the first two, were "less congestion" and "good road design." The order was the same in all 5

a. BOCHUM

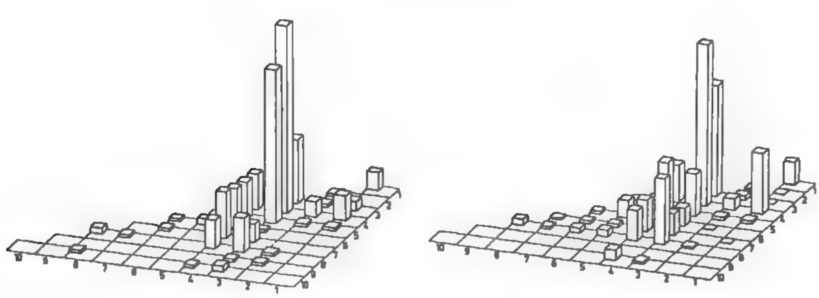
	1	2	3	4	5	6	7	8	9	10
10	—	—	7 / 0.8	4 / 0.4	—	—	—	3 / 0.4		
9	—	4 / 0.4	4 / 0.4		—	—	—	—	—	
8	—	3 / 0.4	40 / 4.5	33 / 3.7			4 / 0.4	7 / 0.8		
7	—	—	22 / 2.4	55 / 6.1		3 / 0.4				
6	—	—	4 / 0.4	48 / 5.3	7 / 0.8	3 / 0.4				
5	4 / 0.4	4 / 0.4	181 / 20.0	44 / 4.9						
4	30 / 3.3	18 / 2.0	221 / 24.5	33 / 3.7						
3	11 / 1.2	4 / 0.4	73 / 8.1							
2		4 / 0.4								
1	22 / 2.4									

b. FRANKFURT

	1	2	3	4	5	6	7	8	9	10
10			7 / 0.3	29 / 1.3						
9										
8	—	4 / 0.2	194 / 8.6	90 / 4.0	20 / 0.9	16 / 0.7	7 / 0.3	18 / 0.8		
7	2 / 0.1		65 / 2.9	92 / 4.1	11 / 0.5	9 / 0.4	4 / 0.2			
6		3 / 0.1	52 / 2.3	70 / 3.1	51 / 2.3	9 / 0.4				
5	8 / 0.4	9 / 0.4	124 / 5.5	141 / 9.6	25 / 1.1					
4	177 / 7.9	34 / 1.5	463 / 17.0	105 / 4.7						
3	2 / 0.1	13 / 0.6	318 / 14.2							
2	2 / 0.1	5 / 0.2								
1	65 / 2.9									

Above: Absolute frequency

Below: Per cent frequency



Columns show relative frequency

Motives for route choice:

1. There exists only one route
2. Route is driven by habit
3. Fastest route
4. Shortest route
5. Good road design
6. More safety by traffic control
7. Right-of-way at most intersections
8. Less congestion
9. Less hilly
10. Other reasons

Figure 2. Route-choice motivation.

TABLE 1

MOTIVATIONS OF DRIVERS WHO HAD THE CHOICE BETWEEN AN URBAN OR RURAL HIGHWAY AND A FREEWAY

Motivation	Motivations of Drivers Who Preferred Urban or Rural Highway When Measured Travel Time and Distance Showed This Route To Be								Motivations of Drivers Who Preferred Freeway When Measured Travel Time and Distance Showed This Route To Be							
	Not Shorter but Faster		Shorter and Faster		Not Shorter and Not Faster		Shorter and Not Faster		Not Shorter but Faster		Shorter and Faster		Not Shorter and Not Faster		Shorter but Not Faster	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
1	9	1.5	23	1.5	5	0.9	9	3.0	8	0.8	11	1.9	1	1.4	12	2.9
2	11	1.9	40	2.6	10	1.8	6	2.0	17	1.7	16	2.8	1	1.4	28	6.7
3	187	32.0	469	30.0	164	29.0	94	31.0	446	45.3	145	25.5	37	52.8	87	20.8
4	157	26.8	468	30.0	139	24.6	96	31.6	111	11.3	218	38.4	2	2.9	174	41.4
5	91	15.6	198	12.6	86	15.2	32	10.5	251	25.4	66	11.6	24	34.3	29	7.0
6	43	7.4	100	6.4	39	6.9	11	3.6	49	5.0	58	10.2	2	2.9	36	8.6
7	43	7.4	123	7.9	47	8.3	23	7.6	52	5.3	44	7.7	1	1.4	30	7.1
8	43	7.4	138	8.8	65	11.5	33	10.8	45	4.6	9	1.6	2	2.9	23	5.5
9	1	0.2	6	0.4	9	1.6			6	0.7	1	0.2			1	0.2
10					1	0.2										
Total	585		1,565		565		304		985		568		70		420	

cities but with some differences in relative importance. The order of the stated motivations and the relative values the drivers attached to them in determining route preference gives a good indication of the existing highway network and traffic situation in these cities. For example, Figure 2 shows that the drivers take advantage of Bochum's good highway system, which offers them fast, short, and well-designed roads, and that drivers in Frankfurt, besides time and distance, prefer routes with low traffic densities and less congestion.

Although travel time and length of a route were those parameters that were evaluated most by drivers, comparison with measured values of travel time and length showed discrepancies in many cases. This indicates that the determination of the best route by drivers is rather subjective. Table 1 gives the comparison and evaluation of statements made by drivers who had the choice of taking an autobahn or freeway route for more than 50 percent of their trip length of origin to destination distances of more than 4 miles versus taking a typical rural or urban highway route with the actually measured values of time and length. The data show that the statements of those drivers who prefer the freeway are much more correct than the statements of the other group with regard to travel time and route length. The judgment of the nonfreeway group about advantages in time or distance on their rural or urban route are not very convincing. For this group, advantages in distance are also improvements in travel time because the difference in travel speeds on alternative routes is very small. Many freeway drivers stated they took the freeway route because of its better design, whereas drivers of the other group declared there was less congestion on their urban route. It may be, however, that the latter really assume they have a better possibility to drive around a point of congestion in front of them if one develops.

STUDY RESULTS

Frequency Distribution on Alternative Routes

For the case of only 2 available routes, the preference and choice of one of these routes by the drivers may range between the 2 extreme possibilities that (a) the assumed resistance of both routes is the same and accordingly there is an equilibrium in the route acceptance, or (b) the assumed route resistance of both differs very strongly and only one of them is accepted. This may be expressed by the following equations:

$$\frac{T_2}{T_1} = 1 \tag{1}$$

$$\frac{T_2}{T_1} = 0 \tag{2}$$

where $T_1 + T_2 = T_{ij}$, the total number of drivers traveling between zones i and j and taking route 1, T_1, or route 2, T_2.

The attractivity or acceptance ratio T_2/T_1 is dependent on several factors that form the route resistance in the driver's mind. Some of them were mentioned previously, but many of them are not measurable objectively. Because the purpose of this study was for application in traffic assignment, only those factors were considered that were measurable. About 30 factors were considered. About 50 different model formulations were tested and evaluated by means of mathematical statistics. The findings were that the most important factors determining or influencing route choice by drivers on trips from home to work are travel time, distance, maximum traffic volume per lane, and number of possible stops. These factors coincide remarkably well with the most mentioned reasons of preference for route choice. Trip diversion can be most effectively described by the following equation:

$$\frac{T_2}{T_1} = (a + b)\left(\frac{t_1}{t_2}\right)^{\alpha} \left(\frac{l_1}{l_2}\right)^{\beta} \left(\frac{q_{max1}}{q_{max2}}\right)^{\gamma} \left(\frac{s_1}{s_2}\right)^{\delta} \tag{3}$$

where

t_1 = total travel time on route 1 (inclusive of any waiting time at the intersection at the first point of choice),

l_1 = total length of route 1,

$q_{max\,1}$ = maximum peak-hour volume per lane that occurs on a road section or link of route 1, and

s_1 = number of possible stops on route 1 (intersections, railroad crossings, or pedestrian crosswalks).

These values are only between the 2 points of choice of alternate routes because tests showed that drivers consider only the gains and advantages of alternative routes between these points of choice. Because Eq. 3 is not a simple equation of proportionality because of its constants a and b, routes 1 and 2 must be defined. Tests showed that route 1 can usually be best defined as the route of shortest travel time. However, for $t_1/t_2 \geq 0.96$ and for $l_1/l_2 \leq 0.80$, route 1 should be defined as the shortest route.

The analysis of the observed values shows that the constants and exponents of the model are dependent on the length of route 1 (Fig. 3). For example, on short distances between the points of choice, the weighting exponents α and β are small because drivers cannot realize the gains in time and length of 1 route versus the other or do not attach much importance to it. However, as the distance becomes larger, both exponents increase; and, after the distance exceeds 4 miles, α is still rising slightly while β starts to decrease after a threshold is reached at 4 miles. The exponent γ is negative for short distances. An analysis of the reported and observed data sets shows that this occurs where the faster route 1 had a good road design with multiple lanes and higher maximum volumes per lane and usually bypassed the city core or fringe areas with high densities, whereas route 2 crossed these areas on roads with lower capacities. Thus, the higher volume q_{max} on route 1 expresses indirectly the attractiveness of this route. For longer distances, γ becomes positive. This means the route with the lower volume is then the more accepted. The exponent δ only exists on short distances and has its implication primarily in high-volume areas. Thus, on short distances the route choice is strongly influenced by the number of possible stops such as at intersections and railroad crossings.

Although the exponents weight the ratios of the factors of resistance that the model contains, the constants a and b can be assumed to consider all other factors of influence that are not in this equation. Because they vary more for short routes than for longer ones, it must be assumed that in this range factors other than t, l, q_{max} and s are important, too.

The coefficient of determination, R^2, for the model was 90.6 percent. The model was also tested without the constants a and b. The run of the exponent curves, however, did not change significantly. The coefficient of determination then was $R^2 = 81.7$ percent, still very good.

For further analysis, the data sets were stratified by the type of traversed town districts (city and core area, suburban and fringe area, and rural area, and freeway use). The parameter curves show typically that the drivers' behavior for their route choices when traversing these areas (Fig. 4) was as described before. The coefficient of determination, however, is much stronger for rural area data ($R^2 = 94.2$ percent) than for core area data ($R^2 = 83.8$ percent). On the core area trips, apparently factors other than those that the model considers have influence on the route choice also. The constants a and b vary considerably in this area.

The model was also applied to cases involving multiple splits, i.e. with 3 and 4 alternative routes, and showed a coefficient of determination $R^2 = 86.7$ percent. These multiple splits were also apportioned into double splits, and only the route sections between the relevant points of choice were considered (1).

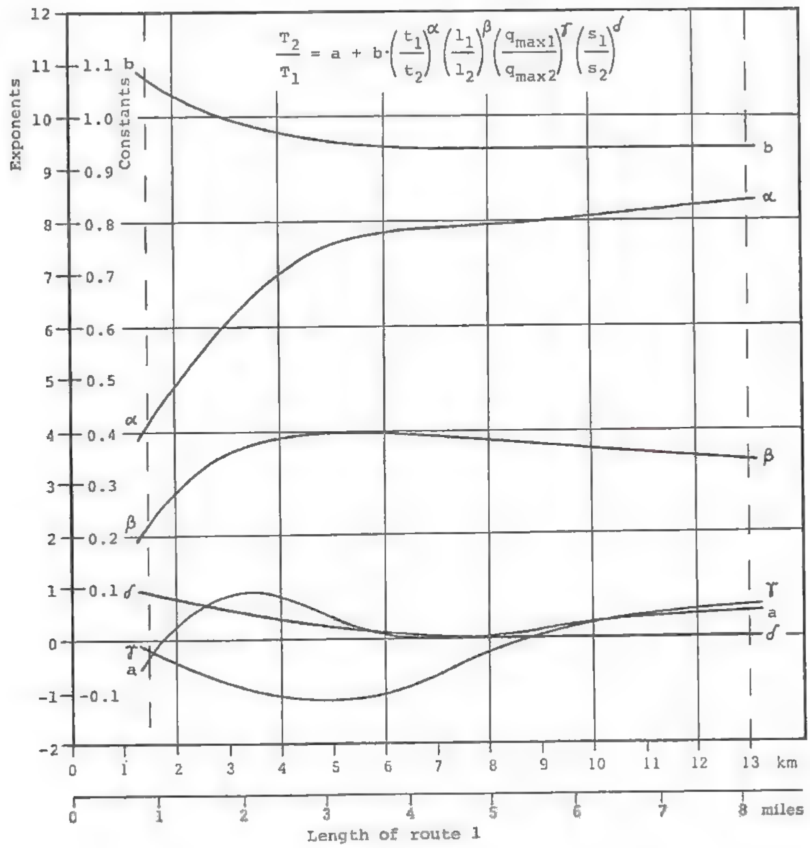


Figure 3. Parameter values for the trip-diversion model.

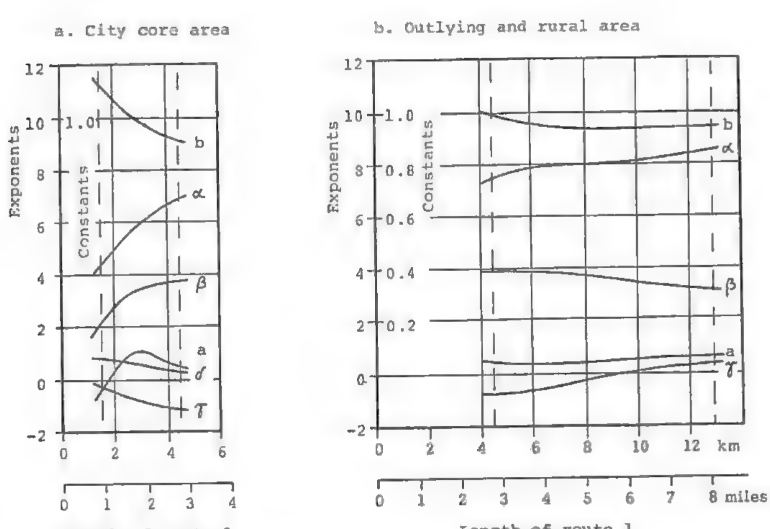


Figure 4. Parameter values for the trip-diversion model by type of urban area.

Evaluation of Alternative Routes

In Frankfurt, for 1 origin and destination interchange, about 20 different routes were found. Most of them, however, were routes of single drivers. Apparently, the number of routes in gridiron networks with roads and intersections of nearly equivalent capacity are similar to the many choices for turn movements available to drivers. However, loops or other illogical routes were never observed. Based only on routes that were taken by more than 3 drivers of the sample from the same origin area, boundary values (minimum and maximum ratios of some factors of route resistance) were developed and show that nonlogical routes will not be accepted by drivers. These values were measured between the points of choice for 2 alternative routes. If the definition of routes 1 and 2, given previously, is used, the observed extreme ratios were $1.0 > t_1/t_2 > 0.62$, $1.6 > l_1/l_2 > 1.58$, $2.4 > q_{max1}/q_{max2} > 0.40$, and $2.3 > s_1/s_2 > 0.29$. One may assume that the range of these ratios also depends on the distance between the points of choice (for short route splits, these ratios will be more extreme than for longer ones), but the investigation showed no significant relationship to the route length. These ratios provide a method of developing logical alternative routes for a traffic assignment program.

The model was applied in a traffic assignment program (with capacity restraint) and compared with other assignment techniques in a transportation study for Muelheim/Ruhr. After 2 iterations, the model showed good results, while other methods needed 6 and more iterations to yield satisfying results (1).

A Proposal for a New Resistance Criterion

There exist some path-finding algorithms to find the best route in a network for any given criterion of route resistance. Most programs use shortest travel time or length as the criterion. However, the routes of shortest time are often those that follow expressways and major street arterials and that overload such routes during the assignment process. The shortest distance criterion, on the other hand, often overloads minor highways. Other studies attempted to eliminate these problems by combining time and length in a new and better criterion of route resistance, for instance by a linear combination $W = a \cdot (\text{travel time}) + b \cdot (\text{travel distance})$. The use of this equation gave results better and more realistic than the use of only time or length gave.

The fact is, as stated before, that there generally are alternative routes between 2 zones because drivers use different criteria for the evaluation of possible routes or they want to take the best route but cannot find it. However, between the 2 zones there is always 1 route that is used by a plurality of all drivers. This route could be used for the definition of the optimum route. Of course, this route may change during the day as the factors determining the criterion of resistance and the trip purposes change.

In his mind, the driver determines the total resistance of a route by weighting and summarizing the factors of resistance on particular sections and points of the highway network in the direction of his destination. Thus, every link on possible paths with its road and traffic characteristics is evaluated by the driver.

The previously mentioned definition for the optimal route was used in making several trials in this research to find a formula for link resistance that minimizes the route resistance of the most frequented route compared with the other less-frequented routes. The result for the resistance of a network link was

$$w_{link} = t^{\alpha} \cdot l^{\beta} \cdot q_{lane}^{\gamma} \qquad (4)$$

where

t = travel time, driving time + waiting time (sec);
l = link length times 3.28 (ft); and
q_{lane} = volume per lane (vph).

The exponents are dependent on the link length. Table 2 gives data showing that drivers attach a greater importance to travel time, length, and volume on short links than on long links because the weighting exponents are then higher.

TABLE 2

PARAMETER VALUES FOR OPTIMAL ROUTE LINK RESISTANCE

Length of Street Section (ft)	α	β	γ
Less than 1,600	0.62	0.40	-0.05
1,600 to 3,200	0.60	0.40	-0.06
More than 3,200	0.56	0.35	-0.07

It was shown that in 95.7 percent of all data sets, the most frequented route was the minimum resistance route if the link resistance was defined as the model. Equation 4 gives the possibility of finding the best path between 2 zones by a computer route-choosing program. However, the collected data showed that travel time itself is a very good criterion for finding the optimum route because in 86.3 percent of all sets, the shortest route in time was the most frequented one. Where the trip length was used as the criterion, the percentage decreased to 65.0 percent; the percentage is higher, however, on short route splits than on longer route splits because distance and travel time then become equivalent.

ACKNOWLEDGMENTS

This research was done with the support of the Ministry of Transportation of the federal government of Germany during 1965-1967. The author is pleased to find some parallels of research in this field in the United States (2), (3). Finally he wants to thank Harold L. Michael for reviewing this article.

REFERENCES

1. Ueberschaer, M. H. Die Aufteilung der Verkehrsstroeme auf verschiedene Fahrtwege (Routen) in Stadtstrassennetzen aufgrund der Strassen- and Verkehrsbedingungen beim morgendlichen Berufspendelverkehr. Strassenbau und Strassenverkehrstechnik, 1969, Heft 85, Bundesanstalt füer Strassenwesen, 5 Köln, Germany.
2. Wachs, M. Relationships Between Drivers' Attitudes Toward Alternate Routes and Driver and Route Characteristics. Highway Research Record 197, 1967, pp. 70-87.
3. Grecco, W. L. Proposed Changes in Urban Transportation Studies. Purdue Univ., Eng. Bull., Series 133, 1969.

PRECISE DETERMINATION OF EQUILIBRIUM IN TRAVEL FORECASTING PROBLEMS USING NUMERICAL OPTIMIZATION TECHNIQUES

Dennis F. Wilkie and Robert G. Stefanek,
Ford Motor Company, Dearborn, Michigan

In recent years, the techniques for planning improvements to transportation systems as well as for analyzing innovative new systems have been given increased attention in the professional literature. Travel forecasting is an important aspect of the planning process because it is necessary to forecast the pattern and magnitudes of traffic flows in the proposed system so that one can analyze the benefits and costs that will accrue to the users and operators of the system. If it is assumed that the number of trips to be made within a region is dependent on the level of service delivered by the transportation system, the problem of determining equilibrium between supply and demand for a given region and transportation system is fundamental to travel forecasting. In this paper, 2 new algorithms are presented for the precise determination of equilibrium in the travel-forecasting problem. A functional of the demand and service variables associated with a transportation system is introduced, and it is shown that the maximum of this functional occurs at equilibrium. Both a constrained gradient and a modified Newton-Raphson algorithm are then used to determine the network flows that maximize this functional, i.e., the equilibrium flows. Two simple examples are considered to demonstrate the use of the algorithms. The advantage of the algorithms presented over present techniques is that equilibrium is obtained precisely rather than approximately and computation of minimum paths is not required in the iterative process.

•IN RECENT YEARS, the techniques for planning improvements to transportation systems as well as for analyzing innovative new systems have been given increased attention in the professional literature (1, 2, 3, 4, 5, 6, 7). Existing techniques have been criticized and new ones proposed (1, 2, 3, 4, 7). Travel forecasting is an important aspect of the planning process. To analyze the benefits and costs that will accrue to the users and operators of a proposed transportation system or to the government and society as a whole requires that a forecast be made of the pattern and magnitudes of traffic flows in the proposed system. The traffic flows in a transportation network result from the interaction between the demand for transportation services in a region and the service characteristics of the transportation system. It has been pointed out by Kraft and Wohl (1), and earlier by Beckman, McGuire, and Winsten (8), that the number of trips that will be made within a region is not independent of the level of service delivered by the transportation system. Thus, the problem of determining the equilibrium between supply and demand for a given region and transportation system is fundamental to the transportation planning process.

For clarity, it is important to briefly review the concepts of equilibrium between supply and demand as applied to transportation networks. The discussion follows

Sponsored by Committee on Transportation Forecasting.

Beckman, McGuire, and Winsten (8). Consider first an isolated transportation link along which the demand for travel as a function of the travel time on the link is as shown in Figure 1a, and the travel time along the link is as shown in Figure 1b. In this case, the equilibrium flow rate and travel time on the link are obtained, by using fundamental microeconomic theory (9), as the intersection of the supply and demand curves as shown in Figure 1c. At this point, the number of trips made, v_e, leads to a travel time on the link t_e, which would in turn imply the number of trips to be made would be v_e.

Generalization of this fundamental idea of equilibrium for an isolated link to a transportation network requires some care. The fundamental assumption on which network equilibrium concepts are based is that all travelers choose to travel along their personal minimum-cost paths through the network from their origins to their destinations. This is Wardrop's first principle (10). It follows that, at equilibrium, if more than 1 path is used by travelers from a given origin to a given destination, the costs along the alternative paths must be the same. Furthermore, the number of trips generated per unit time between all origins and destinations corresponds to the equilibrium network service conditions between the origins and destinations.

One must be careful when trying to mathematically define equilibrium. Beckman, McGuire, and Winsten (8) have precisely defined equilibrium in terms of the demand and service functions associated with a region and its transportation system. However, their proposed technique for determining the network equilibrium essentially ignores their mathematical definition and is approximate. Other approximate techniques for determining network equilibrium are the standard Federal Highway Administration (FHWA) assignment package (11) and the transportation network analysis software package, DODOTRANS, developed at M.I.T. (12). Beckman, McGuire, and Winsten (8), noted the difficulty in getting their iterative approach to converge to an approximate equilibrium, and one can only guess how close to the true equilibrium the approximation would be. Similar convergence problems exist with the FHWA approach. On the other hand, DODOTRANS will always reach an approximation to equilibrium, but again, one can validly ask whether the approximation achieved is even close to true equilibrium.

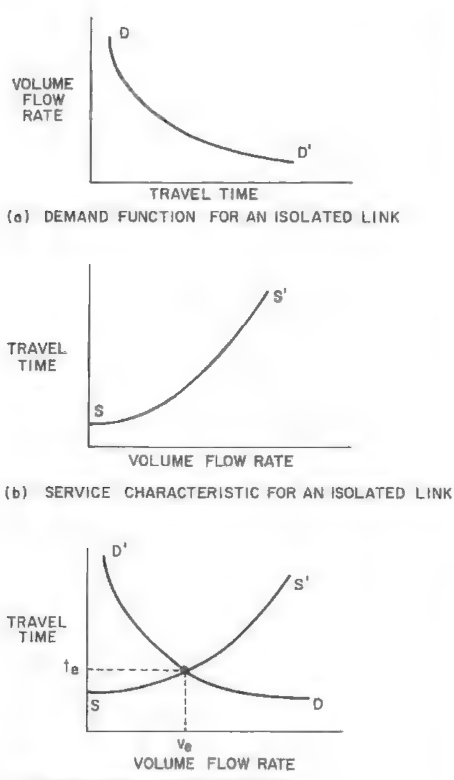


Figure 1. Equilibrium for an isolated link.

In a previous paper (13), numerical techniques of functional maximization were used to find the equilibrium flows and levels of service for a transportation system under the assumption that all interzonal trips are assigned along the unloaded minimum time paths through the network. Because of this assumption, the equilibrium found by that technique differs from the equilibrium defined by Beckman, McGuire, and Winsten (8). In this paper, numerical techniques of functional maximization are used to develop new algorithms for finding equilibrium as defined elsewhere (8).

In the next sections, a mathematical definition of equilibrium is presented based on Beckman's work. A function of the demand and service variables associated with a transportation system is introduced such that the maximum of this function occurs at equilibrium. Algorithms are then developed to maximize this function by iterative adjustment of the network flows. Both a gradient (steepest

descent) method and a Newton-Raphson algorithm are considered, and various examples are presented to exhibit the characteristics of the algorithms.

MATHEMATICAL STATEMENT OF EQUILIBRIUM

The usual means of representing travel in a region is assumed. Namely, the region is partitioned into a set of disjoint zones. The trips from (or to) any zone are assumed to originate (or terminate) at the zone centroid. The transportation network by which interzonal trips are made is represented by a series of links connecting the zones as shown in Figure 2. In general, the links in a network are characterized by a vector of service parameters that may depend on the volume on the link. In the following discussions, the only service variable used to characterize the links is travel time.

Let link ij denote a one-way link directly connecting nodes i and j, and on which persons may travel from node i to node j. (Note that zones and nodes will be used interchangeably through the rest of the paper. There is often a distinction made between a zone being a representation of an area where trips originate or terminate in a region and a node as being a junction of two or more links in a network. This distinction is not made here.) Let $v_{i,j}$ denote the volume flow rate on link ij, and let $v^k_{i,j}$ denote the portion, if any, of the trips $v_{i,j}$ having node k as ultimate destination. Summing the trips on a link over all possible destinations yields the relation

$$v_{i,j} = \sum_k v^k_{i,j} \tag{1}$$

By definition, all $v^k_{i,j}$ and $v_{i,j}$ are non-negative.

Let v^k_i represent the number of trips per unit time originating at node i and having destination k. Then the dependence of the number of trips v^k_i on the level of service in the transportation network is expressed by the demand function

$$v^k_i = d^k_i\ (t^k_i) \text{ if } t^k_i \geq 0 \tag{2}$$

where t^k_i is the travel time through the network from zone i to zone k. The demand function d^k_i is assumed to be a monotone nonincreasing function for $t^k_i \geq 0$. In all subsequent developments, it is also assumed that the inverse of the demand function in Eq. 2 exists; i.e., a function g^k_i exists such that

$$t^k_i = \{d^k_i\}^{-1}\ (v^k_i) = g^k_i\ (v^k_i) \tag{3}$$

The case when g^k_i does not exist is discussed in an example.

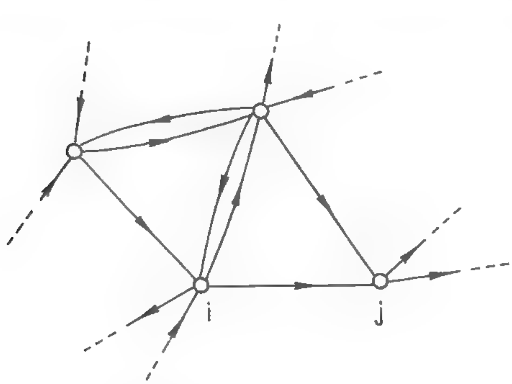


Figure 2. Network representation of a transportation system.

For each node i, let $\{a_i\}$ denote the set of all nodes directly connected to i by a link carrying traffic away from i, and let $\{b_i\}$ denote the set of all nodes directly connected to i by a link carrying traffic toward i. Thus, in Figure 3, $\{a_1\} = \{3,5\}$ $\{b_1\} = \{2,3,4\}$ $\{b_3\} = \{1,2,4\}$, and so forth. Then, the number of trips from node i to node k, v^k_i, obviously equals the flow with destination k away from origin node i minus the flow with destination k toward i. That is,

$$v^k_i = \sum_{\{a_i\}} v^k_{i,a_i} - \sum_{\{b_i\}} v^k_{b_i,i} \tag{4}$$

If $\{a_i\} = \{3,4,7\}$, then define

$$\sum_{\{a_i\}} v^k_{i,a_i} = v^k_{i,3} + v^k_{i,4} + v^k_{i,7}$$

To relate the level of service in the network to the trips using links, requires that supply functions be introduced. Let $t_{i,j}$ denote the time required to travel from node i to node j along the link ij. The relationship between the volume of traffic on the link ij and travel time on the link ij is expressed by the supply function

$$t_{i,j} = h_{i,j}(v_{i,j}) \tag{5}$$

where $h_{i,j}$ is a monotone nondecreasing function.

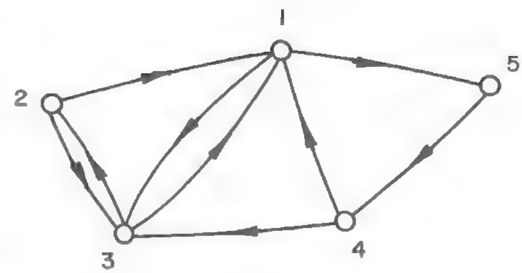


Figure 3. An illustrative network.

As previously explained, the trips from 1 zone to another at equilibrium use a minimum time path between those zones. Hence, $v_{i,j}^k$ is greater than zero at equilibrium if and only if link ij is part of a minimum time path from node i to node k. (Note that there is not necessarily a unique minimum time path between zones at equilibrium.) Therefore, the equilibrium travel time from zone i to zone k can be expressed in terms of the travel time on any link ij for which $v_{i,j}^k > 0$ and the equilibrium travel time from j to k is

$$t_i^k = t_{i,j} + t_j^k \text{ if } v_{i,j}^k > 0 \tag{6}$$

This statement is equivalent to the well-known principle of optimality stated by Bellman (14).

On the other hand, if $v_{i,j}^k = 0$, link ij is not part of a minimum time path from i to k. It follows that the equilibrium travel time from j to k plus the equilibrium travel time on link ij must be greater than (or possibly equal to) the time to travel from i to k at equilibrium along a path that does not include link ij; that is,

$$t_j^k + t_{i,j} \geq t_i^k \text{ if } v_{i,j}^k = 0 \tag{7}$$

Equations 6 and 7 give a precise mathematical statement of equilibrium.

Substitution of the inverse demand functions from Eq. 3 and the service functions from Eq. 5 into Eqs. 6 and 7 yields the following alternate statement of equilibrium in terms of the network flows:

$$g_i^k(v_i^k) - g_j^k(v_j^k) = h_{i,j}(v_{i,j}) \text{ if } v_{i,j}^k > 0 \tag{8}$$

$$g_i^k(v_i^k) - g_j^k(v_j^k) \leq h_{i,j}(v_{i,j}) \text{ if } v_{i,j}^k = 0 \tag{9}$$

Thus, the problem of finding equilibrium in a travel-forecasting problem is equivalent to determining the values of the non-negative flows $v_{i,j}^k$ that satisfy Eqs. 1, 4, 8 and 9. However, determining the solution of these simultaneous nonlinear equations, without knowing what the minimum time paths between zones are at equilibrium, is a difficult task and has led to the use of approximate iterative techniques to determine equilibrium.

In the next section, the statement of equilibrium as the maximum of a functional of the network flows is developed, following Beckmann, McGuire, and Winsten (8).

EQUILIBRIUM AS A MAXIMIZATION PROBLEM

For convenience, define a vector $\underline{v}$ that has all $v_{i,j}^k$ as its components; i.e., $\underline{v}$ transpose is given by

$$\underline{v}^T = [v_{1,j}^2, \ldots v_{i,j}^k, \ldots] \tag{10}$$

Consider the functional

$$H(\underline{v}) = \sum_m \sum_\ell \int_0^{v_m^\ell} g_m^\ell(x)\,dx - \sum_m \sum_n \int_0^{v_{m,n}} h_{m,n}(x)\,dx \tag{11}$$

where each summation is over all node numbers. It is shown in the Appendix that the gradient components of $H(\underline{v})$, $\partial H/\partial v_{i,j}^k$, are given by

$$\frac{\partial H}{\partial v_{i,j}^k} = g_i^k(v_i^k) - g_j^k(v_j^k) - h_{i,j}(v_{i,j}) \tag{12}$$

Now, Eq. 8 and Eq. 12 together imply that, at equilibrium,

$$\frac{\partial H}{\partial v_{i,j}^k} = 0 \text{ if } v_{i,j}^k > 0 \tag{13}$$

whereas Eq. 9 and Eq. 12 imply that at equilibrium

$$\frac{\partial H}{\partial v_{i,j}^k} \leq 0 \text{ if } v_{i,j}^k = 0 \tag{14}$$

Equations 13 and 14 are by definition the necessary conditions for the functional H to have a maximum at a point $\underline{v}$ under the constraint that all $v_{i,j}^k$ be non-negative. The fact that Eqs. 13 and 14 are satisfied at equilibrium implies that finding equilibrium is equivalent to the problem of maximizing H subject to the constraint that $v_{i,j}^k \geq 0$ for all i, j, k. It is shown in the work by Beckman, McGuire, and Winsten (8) that the solution to the problem of maximizing H is unique (in the sense that the equilibrium flows $v_{i,j}$ on all roads are unique) whenever the inverse demand functions g_i^k are strictly decreasing functions of v_i^k and the supply functions $h_{i,j}$ are strictly increasing functions of $v_{i,j}$.

The question is, Does the statement of equilibrium as a functional maximization make the problem of finding equilibrium easier to solve? It is the authors' opinion that the answer to this question is yes, because numerical techniques of functional maximization can readily be used to determine the equilibrium network flows that maximize $H(\underline{v})$. Such numerical techniques are used effectively in parameter optimization problems associated with control system design (15), and it is shown in the following sections that these techniques can be used in the solution of the problem being considered.

ALGORITHMS FOR DETERMINING EQUILIBRIUM

Two algorithms that use iterative numerical techniques to find the values $v_{i,j}^k$, $\forall$ i, j, k, which maximize H (i.e., to find equilibrium), are presented in this section. Both algorithms utilize the vector $\nabla_v H$, the gradient of H with respect to $\underline{v}$. By definition

$$\nabla_{\underline{v}} H^T = \left[\frac{\partial H}{\partial v_{1,j}^2} \cdots \frac{\partial H}{\partial v_{i,j}^k} \cdots \right] \tag{15}$$

and thus the components of $\nabla_v H$ are obtained by using Eq. 12.

The algorithms presented are a constrained gradient algorithm and a modified Newton-Raphson procedure. The rationale for these algorithms is not discussed here but can be found in any standard reference for optimization techniques (16).

In the constrained gradient algorithm, the change in each volume at the end of the r th iteration is given by

$$\Delta v_{i,j}^{k}\Big|_{r} = v_{i,j}^{k}\Big|_{r+1} - v_{i,j}^{k}\Big|_{r} = \begin{cases} 0 & \text{if } v_{i,j}^{k}\Big|_{r} = 0 \text{ and } \dfrac{\partial H}{\partial v_{i,j}^{k}}\Big|_{r} < 0 \\ \alpha \dfrac{\partial H}{\partial v_{i,j}^{k}}\Big|_{r} & \text{if otherwise} \end{cases} \tag{16}$$

In Eq. 16, α is a positive scalar constant that is chosen to guarantee that none of the flows is adjusted to be negative or to locally maximize the functional H in the direction defined by the gradient. A flow chart for the algorithm is given in Figure 4.

The iteration procedure continues until at some step Eqs. 13 and 14 are satisfied to within the desired degree of accuracy. Typically, this means that iteration continues until

$$-\epsilon < \frac{\partial H}{\partial v_{i,j}^{k}}\Big|_{r} < \epsilon \tag{17}$$

for all $v_{i,j}^{k}|_r > 0$, where ϵ is a small scalar constant (e.g., 10^{-3}). Results using this algorithm to determine equilibrium are given in the examples.

A second approach to finding the non-negative $v_{i,j}^{k}$'s that maximize H is to use a modified Newton-Raphson algorithm. If the $v_{i,j}^{k}$'s were not constrained to be non-negative, then the change in the volumes at the end of the r th iteration, $\Delta v|_r$, using the Newton-Raphson procedure would be given by

$$\Delta \underline{v}\big|_{r} = -\left[\frac{\partial \nabla_{\underline{v}} H}{\partial \underline{v}}\Big|_{r}\right]^{-1} \nabla_{\underline{v}} H\big|_{r} \tag{18}$$

(Thus, the second partial derivatives of H with respect to all $v_{i,j}^{k}$'s must be computed and the matrix of these terms inverted in the Newton-Raphson procedure.) However, because of the constraint that all of the components of $\underline{v}$ must be non-negative, a modified Newton-Raphson algorithm must be used in the problem being considered.

Basically, the idea of the algorithm is to eliminate from consideration, at step r of the iteration procedure, all components of $\underline{v}$ that satisfy Eq. 14 at step r. Therefore, at step r, determine the number c of components of $\underline{v}|_r$ that do not satisfy Eq. 14. Define $\underline{w}^r$ as a vector of dimension c and denote its value at step r as $\underline{w}^r|_r$. (The dimension of $\underline{w}^r$ may be different at each step.) Every component of $\underline{v}|_r$ that does not satisfy Eq. 14 is included sequentially as a component of $\underline{w}^r|_r$.

Then, similar to the iteration procedure given in Eq. 18, a vector $\Delta\underline{w}^r|_r$ is calculated by

$$\Delta \underline{w}^{r}\Big|_{r} = -\left[\frac{\partial \nabla_{\underline{w}^r} H}{\partial \underline{w}^{r}}\Big|_{r}\right]^{-1} \nabla_{\underline{w}^r} H\Big|_{r} \tag{19}$$

Then, the components of the c-dimensional vector $\underline{w}^r|_{r+1}$ are computed by the relations

$$w_{\ell}^{r}\Big|_{r+1} = \max\left\{w_{\ell}^{r}\Big|_{r} + \Delta w_{\ell}^{r}\Big|_{r}, 0\right\},\ \ell = 1, \ldots, c \tag{20}$$

Finally, the components of $\underline{v}|_{r+1}$ are obtained in the following way: The components of $\underline{v}$ that at step r are not members of $\underline{w}^r$ are set to zero in $\underline{v}|_{r+1}$. The remaining

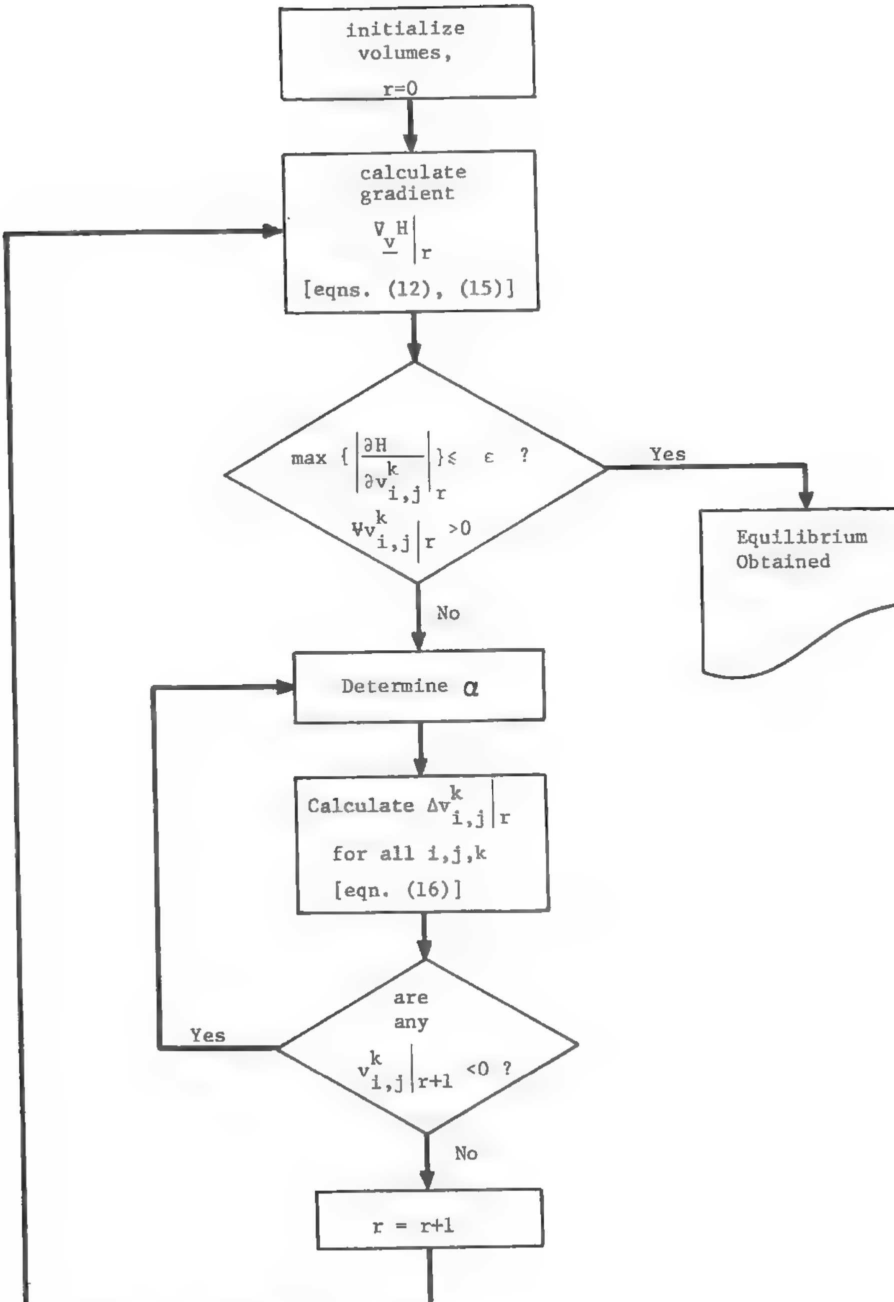


Figure 4. Flow chart for constrained gradient algorithm.

components of $\underline{v}\,|_{r+1}$ are obtained sequentially from $\underline{w}^r|_{r+1}$. The iteration procedure continues until Eq. 17 is satisfied. Figure 5 shows a flow chart for the modified Newton-Raphson algorithm.

The algorithm just presented requires the calculation of $\partial \nabla_{\underline{w}^r} H / \partial \underline{w}^r$, the matrix of second partial derivatives of H with respect to the $v^k_{i,j}$'s that are components of $\underline{w}^r$. Instead of an expression for this matrix being obtained separately for each iteration

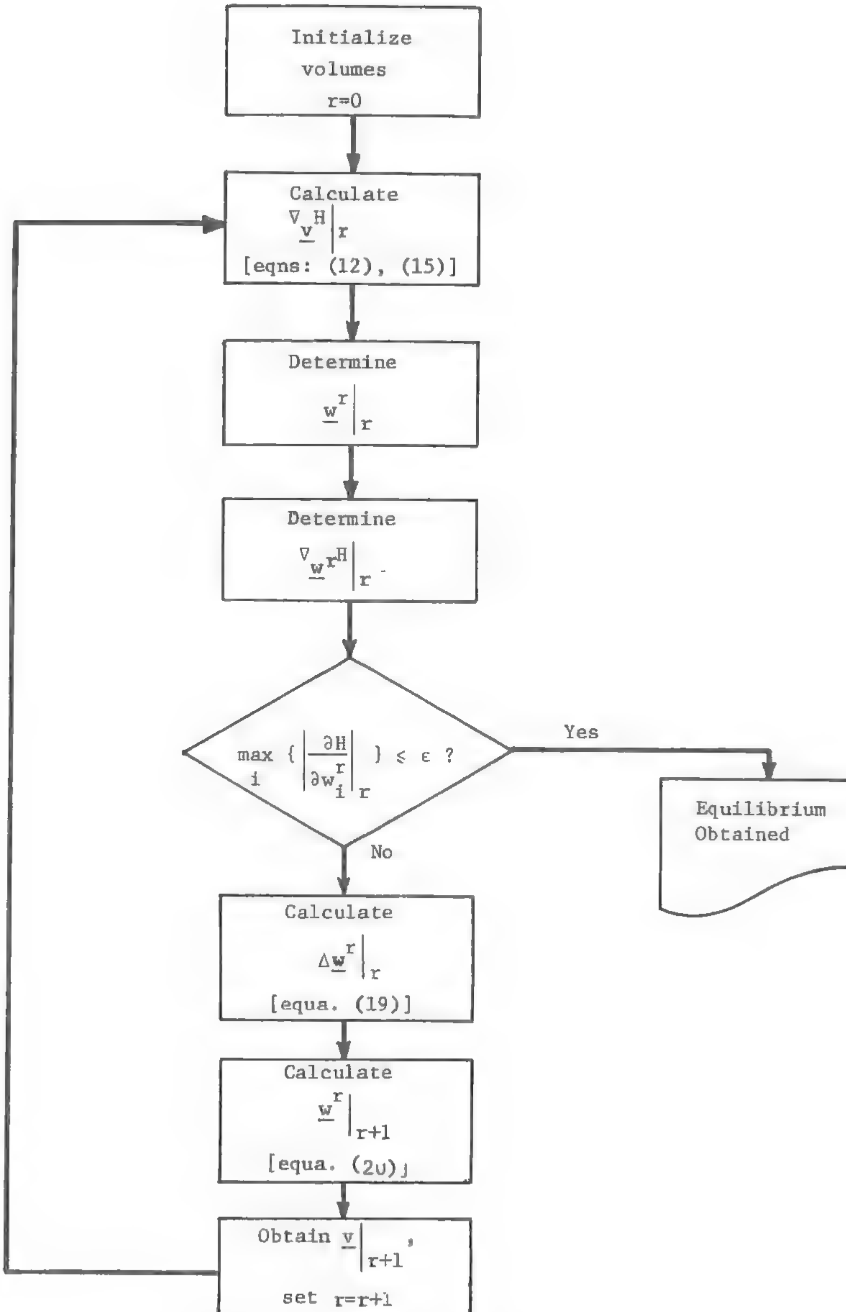


Figure 5. Flow chart for modified Newton-Raphson algorithm.

step, it is easier to obtain an expression for $\partial\nabla_{\underline{v}}H/\partial\underline{v}$, the matrix of second partial derivatives of H with respect to all of the $v^k_{i,j}$'s, and then to recognize that $\partial\nabla_{\underline{w}^r}H/\partial\underline{w}^r$ consists of some of the rows and columns of $\partial\nabla_{\underline{v}}H/\partial\underline{v}$. In addition, it can be shown that

$$\frac{\partial^2 H}{\partial v^{\ell}_{m,n}\partial v^k_{i,j}} = -\dot{h}_{i,j}(v_{i,j})\,\delta_{im}\,\delta_{jn} + \dot{g}^k_i(v^k_i)\,[\delta_{im}\sum_{\{a_i\}}\delta_{n,a_i} - \delta_{in}\sum_{\{b_i\}}\delta_{m,b_i}]\,\delta_{kl} - \dot{g}^k_j(v^k_j)\,[\delta_{jm}\sum_{\{a_i\}}\delta_{n,a_j} - \delta_{jn}\sum_{\{b_j\}}\delta_{m,b_j}]\,\delta_{kl} \tag{21}$$

In Eq. 21, the dot above a function implies differentiation of that function with respect to its argument, $\delta_{ij} = 1$ if i = j and 0 if $i \neq j$; and $\sum_{\{a_i\}}\delta_{n,a_i} \equiv 1$ if n is a member of $\{a_i\}$ and 0 if otherwise.

Thus, all the elements of $\partial\nabla_{\underline{v}}H/\partial\underline{v}$ are available from Eq. 21. Furthermore, if all the inverse demand and supply functions g^k_i and $h_{i,j}$ respectively are linear functions of their arguments, then each component of $\partial\nabla_{\underline{v}}H/\partial\underline{v}$ is a constant.

EXAMPLES

Two relatively simple examples are considered in this section to illustrate the application of the preceding results to equilibrium problems in travel forecasting. The network used in the examples is shown in Figure 6. For simplicity, the demand and service functions in the examples are assumed to be linear. However, this restriction is certainly not necessary, and any nonlinear (or piece-wise linear) functions can readily be used in the preceding algorithms.

Example 1

The interzonal demand and link service functions used in this example are as follows:

$$v^2_1 = d^2_1(t^2_1) = 16.625 - t^2_1 \qquad v^3_1 = 28.25 - t^3_1$$

$$v^4_1 = 21.375 - t^4_1 \qquad v^3_2 = 21.625 - t^3_2$$

$$v^3_4 = 16.875 - t^3_4 \qquad v^3_5 = 28 - t^3_5$$

and

$$t_{1,2} = h_{1,2}(v_{1,2}) = 5 + 0.1\,v_{1,2} \qquad t_{2,3} = 10 + 0.1\,v_{2,3}$$

$$t_{1,4} = 10 + 0.1\,v_{1,4} \qquad t_{4,3} = 5.5 + 0.1\,v_{4,3}$$

$$t_{5,1} = 1 \qquad t_{5,3} = 18$$

Two points to note are that the inverse of every demand functions exists and two of the service functions are independent of the volume flow rate of trips using the link. The relationships between the various volume components can be clarified by considering

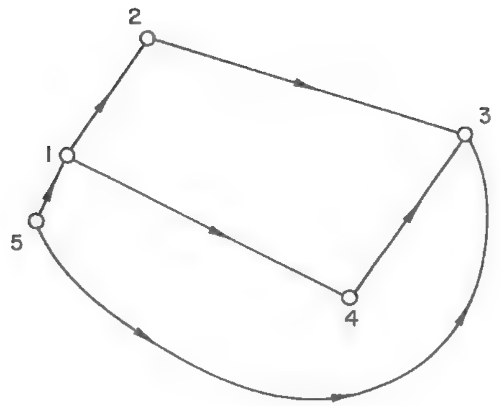


Figure 6. Network for the examples.

the calculation of the gradient component $\partial H/\partial v^3_{1,2}$. From Eq. 12,

$$\frac{\partial H}{\partial v^3_{1,2}} = g^3_1(v^3_1) - g^3_2(v^3_2) - h_{1,2}(v_{1,2}) \tag{24}$$

Inverting the appropriate demand functions given in the preceding yields

$$g^3_1(v^3_1) = 28.25 - v^3_1 \tag{25}$$

$$g^3_2(v^3_2) = 21.625 - v^3_2$$

However, v^3_1 and v^3_2 must be related to the component flows on the various links. Thus, using Eq. 4, we obtain

$$v^3_1 = v^3_{1,2} + v^3_{1,4} - v^3_{5,1} \tag{26}$$

$$v^3_2 = v^3_{2,3} - v^3_{1,2}$$

If we substitute Eqs. 25 and 26 and the appropriate service function in Eq. 24, it follows that

$$\frac{\partial H}{\partial v^3_{1,2}} = 1.625 - 2.1\, v^3_{1,2} - v^3_{1,4} - 0.1\, v^2_{1,2} + v^3_{2,3} + v^3_{5,1} \tag{27}$$

Similar expressions are obtained for all the gradient components $\partial H/\partial v^k_{i,j}$ for use in both the gradient and modified Newton-Raphson algorithm.

Obviously, some initial estimates of the component flows on the various links must be made to use the algorithms discussed earlier. There are various approaches that can be taken in making initial estimates; for example, completely random volumes could be chosen, all zero volumes could be chosen, or volumes corresponding to the unloaded minimum time path could be chosen. The latter 2 cases were tried here. The results for both the gradient and modified Newton-Raphson methods are shown in Table 1. Two sets of stopping criteria, i.e., $\epsilon = 10^{-2}$ and 10^{-3}, were used for the gradient method. The stopping criterion for the modified Newton-Raphson method was $\epsilon = 10^{-7}$. It is seen that both techniques converged to the correct equilibrium. The modified Newton-Raphson method converged in 1 iteration with the zero initial estimates and in 3 iterations with the other estimates. In the latter case, the number of variables considered at each iteration was different. Thus, the modifications in the Newton-Raphson method required by the constraints were indeed exercised. It is interesting to note that

TABLE 1

RESULTS FOR EXAMPLE 1

Function	Initial Guess	Gradient $\epsilon = 10^{-2}$ (196 iterations)	Gradient $\epsilon = 10^{-3}$ (312 iterations)	Initial Guess	Gradient $\epsilon = 10^{-2}$ (135 iterations)	Gradient $\epsilon = 10^{-3}$ (164 iterations)	Modified Newton-Raphson (1 iteration)
$v^2_{1,2}$	11.625	9.989	9.999	0.0	9.998	10.000	10.00
$v^3_{1,2}$	25.25	6.354	6.261	0.0	6.251	6.247	6.25
$v^3_{1,4}$	0.0	3.646	3.739	0.0	3.76	3.753	3.75
$v^4_{1,4}$	11.375	10.010	10.001	0.0	9.998	9.999	10.00
$v^3_{2,3}$	36.875	16.353	16.260	0.0	16.259	16.248	16.25
$v^3_{4,3}$	11.375	13.646	13.739	0.0	13.768	13.753	13.75
$v^3_{5,1}$	12.0	0.0	0.0	0.0	0.0	0.0	0.0
$v^3_{5,3}$	0.0	9.999	9.999	0.0	9.999	9.000	10.0

both the gradient and modified Newton-Raphson algorithms converged more quickly with the zero initial conditions. It is also interesting to note that the unloaded minimum time path from zone 5 to zone 3 is from $5 \rightarrow 1 \rightarrow 2 \rightarrow 3$, whereas the minimum time path from 5 to 3 at equilibrium is along the link directly connecting zones 5 and 3. Thus, the first set of initial estimates assigns all the volume from zones 5 to 3 along the wrong path, but the techniques still converge. Perhaps, however, this explains the need for fewer iterations with zero initial estimates for all link volumes. These results are indeed quite good, especially those using the modified Newton-Raphson algorithm.

Example 2

In this example, the case where the inverses of all demand functions do not exist is considered. This corresponds to a fixed demand between zones. This case is of interest because some of the nodes used in representing transportation networks may be "throughpass" nodes where no trips originate or terminate but where alternate routes intersect.

The technique used here is to define a pseudo-demand function for which an inverse function does exist but which is a good approximation to the fixed, zero demand function. The algorithms presented in this paper can then be applied; the question is simply, Will the algorithms converge?

The network shown in Figure 6 is used again. The volumes v^3_2 and v^4_1 are assumed to be zero, and the demand and service functions are as follows:

$$v^2_1 = d^2_1(t^2_1) = 16.25 - t^2_1 \qquad v^3_1 = 19.0 - t^3_1$$

$$v^3_4 = 16.5 - t^3_4 \qquad v^3_5 = 22.5 - t^3_5$$

The zero volume v^3_2 is approximated by

$$v^3_2 = 10^{-3} - 10^{-5}\ t^3_2$$

and

$$t_{1,2} = h_1^2(v_1^2) = 5 + 0.1\ v_{1,2}$$

$$t_{2,3} = 10 + 0.1\ v_{2,3} \qquad t_{1,4} = 10 + 0.1\ v_{1,4}$$

$$t_{4,3} = 5.5 + 0.1\ v_{4,3} \qquad t_{5,1} = 1 + 0.1\ v_{5,1}$$

$$t_{5,3} = 17 + 0.1\ v_{5,3}$$

No pseudo-demand function was defined for v_1^4 because it turns out that it does not enter into any of the gradient components and can thus be completely ignored. The gradient components are expressed in terms of all the component flows $v_{i,j}^k$ as discussed in the preceding example. One of the gradient components that involves a pseudo-demand function is $\partial H/\partial v_{1,2}^3$, and this is found to be

$$\frac{\partial H}{\partial v_{1,2}^3} = -86 - 1.1\ v_{1,2}^3 - v_{1,4}^3 + v_{5,1}^3 + 10^5\ (v_{2,3}^3 - v_{1,2}^3) - 0.1\ v_{1,2}^2 \tag{28}$$

One would suspect that numerical difficulties would occur when gradient components such as those in Eq. 28 are used. Indeed, the gradient algorithm did not converge for any set of initial estimates tried and for up to 5,000 iterations. However, the modified Newton-Raphson algorithm converged to the equilibrium in two iterations ($\epsilon = 10^{-4}$). This is interesting because the matrix of second partial derivatives that must be inverted contains terms that differ by 5 orders of magnitude. The equilibrium flows in this case are

$$v_{1,2}^2 = 10 \qquad v_{1,2}^3 = 2.5 \qquad v_{1,4}^3 = 0$$

$$v_{2,3}^3 = 2.5 \qquad v_{4,3}^3 = 10 \qquad v_{5,3}^3 = 5 \qquad v_{5,1}^3 = 0$$

SUMMARY AND CONCLUSIONS

The problem of precisely determining the equilibrium between supply and demand was considered in a travel-forecasting problem when travel demand is assumed to depend on the level of service delivered by a transportation system. Based on some results of Beckman, McGuire, and Winsten (8), 2 new algorithms for precisely determining equilibrium were developed by using numerical optimization techniques. The algorithms use a constrained gradient technique and a modified Newton-Raphson procedure to maximize a functional of network flows, and Beckman, McGuire, and Winsten (8) have shown that the flows thus determined are the equilibrium flows in the network. Very good results were obtained in the examples considered by using the algorithms given.

The advantages of the algorithms presented in the paper over techniques like DODOTRANS (12) are that equilibrium is obtained precisely rather than approximately, and computation of minimum paths during the iterative process is not required. Certainly more work remains to be done in applying the techniques to very large networks. However, one can be very optimistic about the application of these techniques to large networks because of known techniques in other fields.

REFERENCES

1. Kraft, G., and Wohl, M. New Directions for Passenger Demand Analysis and Forecasting. Transportation Research, Vol. 1, 1967, pp. 205-230.
2. Wohl, M. Another View of Transport System Analysis. Proc. I.E.E.E., April, 1968, pp. 446-457.
3. Hillegass, T. J. Urban Transportation Planning—A Question of Emphasis. Traffic Engineering, Vol. 19, No. 7, June 1969, pp. 46-48.
4. Heggie, I. G. Are Gravity and Interactance Models a Valid Technique for Planning Regional Transport Facilities? Operations Research Quarterly, Vol. 20, No. 1, 1969, pp. 93-110.

5. Wilson, A. G. A Statistical Theory of Spatial Distribution Models. Transportation Research, Vol. 1, 1967, pp. 253-269.
6. McLynn, J. M., and Watkins, R. H. Multimode Assignment Models. Northeast Corridor Transportation Project, Technical Paper 7, U.S. Department of Commerce, Feb. 1967.
7. Quandt, R. E. Estimation of Modal Splits. Transportation Research, Vol. 2, 1968, pp. 41-50.
8. Beckman, M., McGuire, C. B., and Winsten, C. B. Studies in the Economics of Transportation. Yale Univ. Press, New Haven, Conn., 1956.
9. Henderson, J. M., and Quandt, R. E. Microeconomic Theory. McGraw-Hill, New York, N.Y., 1958.
10. Wardrop, J. G. Some Theoretical Aspects of Road Traffic Research. Proc. Institute of Civil Engineers, Part 2, Vol. 1, 1952, pp. 352-378.
11. Traffic Assignment Manual. U.S. Department of Commerce, Bureau of Public Roads, Govt. Print. Off., 1964.
12. Ruiter, E. R. Techniques for Searching Out Transportation System Alternatives. Summary Report. Transportation Systems Div., Dept. of Civil Eng., Massachusetts Institute of Technology, July 1969.
13. Wilkie, D. F. The Application of Numerical Optimization Techniques to a Travel Forecasting Problem. Paper presented at the Third Hawaii Intern. Conf. on System Sciences, Honolulu, Jan. 14-16, 1970.
14. Bellman, R. Dynamic Programming. Princeton Univ. Press, N.J., 1957.
15. Wilkie, D. F. Application of Some Control Theory Techniques to Transportation Systems Analysis. Transportation Research and Planning Office, Ford Motor Company, Rept. 69-3, Dearborn, Mich., May 1969.
16. Wilde, D. J., and Beightler, C. S. Foundations of Optimization. Prentice-Hall, Englewood Cliffs, N.J., 1967.

APPENDIX

The functional H is given in Eq. 11. If the chain rule for differentiation is used, it follows that

$$\frac{\partial H}{\partial v^k_{i,j}} = \sum_m \sum_\ell \frac{d}{dv^\ell_m}\left[\int_0^{v^\ell_m} g^\ell_m(x)\,dx\right]\frac{\partial v^\ell_m}{\partial v^k_{i,j}} - \sum_m \sum_n \frac{d}{dv_{m,n}}\left[\int_0^{v_{m,n}} h_{m,n}(x)\,dx\right]\frac{\partial v_{m,n}}{\partial v^k_{i,j}} \tag{29}$$

Then, using Leibnitz's rule in Eq. 29 yields

$$\frac{\partial H}{\partial v^k_{i,j}} = \sum_m \sum_\ell g^\ell_m(v^\ell_m)\frac{\partial v^\ell_m}{\partial v^k_{i,j}} - \sum_m \sum_n h_{m,n}(v_{m,n})\frac{\partial v_{m,n}}{\partial v^k_{i,j}} \tag{30}$$

Now, if v^ℓ_m is expressed in the form of Eq. 4, it can be seen that the partial derivative of v^ℓ_m with respect to $v^k_{i,j}$ is zero unless $\ell = k$. Similarly, if $v_{m,n}$ is expressed in the form of Eq. 1, $\frac{\partial v_{m,n}}{\partial v^k_{i,j}}$ is seen to be zero unless $m = i$ and $n = j$. Hence, Eq. 30 becomes

$$\frac{\partial H}{\partial v^k_{i,j}} = \sum_m g^k_m(v^k_m)\frac{\partial v^k_m}{\partial v^k_{i,j}} - h_{i,j}(v_{i,j})\frac{\partial v_{i,j}}{\partial v^k_{i,j}} \tag{31}$$

Again making use of Eq. 4, we obtain

$$\sum_m g_m^k(v_m^k)\frac{\partial v_m^k}{\partial v_{i,j}^k} = \sum_m g_m^k(v_m^k)\frac{\partial}{\partial v_{i,j}^k}\left[\sum_{\{a_m\}} v_{m,a_m}^k\right] - \sum_m g_m^k(v_m^k)\frac{\partial}{\partial v_{i,j}^k}\left[\sum_{\{b_m\}} v_{b_m,n}^k\right] \tag{32}$$

$$g_i^k(v_i^k)\frac{\partial}{\partial v_{i,j}^k}\left[\sum_{\{a_i\}} v_{i,a_i}^k\right] - g_j^k(v_j^k)\frac{\partial}{\partial v_{i,j}^k}\left[\sum_{\{b_j\}} v_{b_j,j}^k\right] \tag{33}$$

Equation 33 was obtained by recognizing that the partial derivatives involved in the first and second terms on the right of Eq. 32 are respectively equal to zero when $m \neq n$ and $m \neq j$. The fact that $v_{i,j}^k$ exists as a variable implies that link ij exists to carry traffic from i to j. By definition, then, one of the members of $\{a_i\}$ is j, and one of the members of $\{b_j\}$ is i. Hence, $v_{i,j}^k$ appears once in each of the summations, $\sum_{\{a_i\}} v_{i,a_i}^k$ and $\sum_{\{b_j\}} v_{b_j,j}^k$; and the partial derivative with respect to $v_{i,j}^k$ of each of these summations thus equals 1. Equation 33 then becomes

$$\sum_m g_m^k(v_m^k)\frac{\partial v_m^k}{\partial v_{i,j}^k} = g_i^k(v_j^k) \tag{34}$$

Using Eq. 1, we have

$$\frac{\partial v_{i,j}}{\partial v_{i,j}^k} = 1 \tag{35}$$

so that substitution of Eq. 34 into Eq. 31 and use of Eq. 35 finally yields

$$\frac{\partial H}{\partial v_{i,j}^k} = g_i^k(v_i^k) - g_j^k(v_j^k) - h_{i,j}(v_{i,j}) \tag{36}$$

CONSIDERATION OF INTERMODAL COMPETITION IN THE FORECASTING OF NATIONAL INTERCITY TRAVEL

Raymond H. Ellis, Paul R. Rassam, and John C. Bennett,
Peat, Marwick, Mitchell and Company, Washington, D.C.

Forecasting of national travel demands has been performed to date separately for each mode, without giving consideration to changes in the transportation services and demands of other modes. A study was recently completed that had as its objectives the development of a prototype methodology for estimating future national passenger travel demands and the exercising of this methodology. This paper describes selected innovative aspects of an overall national intercity travel demand forecasting framework developed in this study. A method is described for transforming air trip tables, which are on an airport-to-airport basis as developed by the Civil Aeronautics Board into trip tables reflecting the "true" origins and destinations of the air trips. Also described is the development of a modal-split model that is based on a logistic function and that estimates the automobile and air market shares as a function of the travel impedances of each of the modes.

●THE IMPORTANCE of considering the competition among modes has long been recognized in travel demand analyses conducted at the urban and corridor levels. Forecasting of national intercity travel has generally been performed separately for each mode, without giving consideration to changes in the transportation services and demands of other modes. Changes in the level of service of a mode, such as a highway, will affect the demands not just for that mode but for the other competing modes, such as rail or air. In this sense, it is important to develop demand forecasts for national intercity travel based on a methodology that explicitly considers the competition among modes of transportation.

A study (1) was recently completed that had 2 main objectives: (a) the development of a prototype methodology for forecasting national intercity travel such that the competition among modes is explicitly introduced into the analysis, and (b) the exercising of this methodology to prepare forecasts for 1975, 1980, and 1990. Although the results of this study are plausible, their interpretation and subsequent use should be considered within the context of the relatively limited resources with which they were developed. This paper focuses on selected innovative highlights of the forecasting framework. Specifically, it describes (a) a method for transforming airport-to-airport trip tables into trip tables reflecting the "true" origins and destinations of the air trips; and (b) a modal-split model that is based on a logistic curve and that estimates the automobile and air market shares as a function of the travel impedances of each of the modes.

Examination of the 1967 National Travel Survey strongly suggests that the predominant amount of intercity travel in the United States involves the automobile and air modes. On a national basis, the rail and bus modes account for approximately 4 percent of the total number of person trips. However, rail trips would tend to be concentrated in certain travel corridors, such as the Northeast Corridor, where they might capture a significant proportion of the trips, particularly if Metroliner types of services

Sponsored by Committee on Transportation Forecasting.

were introduced. For this reason, all potential transportation corridors in the country were identified, and separate demand analyses developed for the corridor and noncorridor trips.

For noncorridor trips, the analysis was restricted to the competition among the automobile, air, and bus modes. The analysis framework described in this paper was designed to predict travel demands, primarily by automobile and air modes, for noncorridor travel. Bus travel demands, not considered for modal-split calibration procedures because of insufficient data, were estimated by using factors obtained from the 1967 National Travel Survey. The reader is referred to the final report (1) for a discussion of the analysis process for corridor trips.

In this study, the areal system consists of 490 zones that are aggregates of counties. In this 490-zone structure, each standard metropolitan statistical area (SMSA) in the continental United States is represented by a single zone; the remainder of each state is divided into a geometrically suitable number of zones.

DEVELOPMENT OF TRUE AIR TRIP TABLES

The basic air travel data source of this study was the Civil Aeronautics Board (CAB) 1967 survey (2) that included a 10 percent sample of (a) all tickets sold by the certificated trunk and local carriers for domestic travel in scheduled services in the continental United States, and (b) all tickets lifted by these carriers for domestic travel and originally issued by airlines not reporting traffic for the CAB survey. Sample design and reliability and collection and reporting instructions are discussed elsewhere (2). However, it is of interest to note that 519 origins or destinations are identified by city codes at the air-hub level, which in many cases consists of more than 1 airport.

Several programs were written that processed the summary CAB tapes and developed a 490-zone origin and destination air trip table. A total of 88,074,230 expanded trips were contained in this table from which 338,910 trips were discarded because of invalid codes. This control total compares favorably with the CAB control total of 88,434,820 trips (2). However, detailed examination of this table revealed a major lacuna: 141 of the 490 zones were not associated with a city code, i.e., an air hub, thus implying that these zones did not attract or generate air trips. The procedure for solving this problem is described later.

The trip tables developed from the CAB sample are at the air-hub level; hence, they do not necessarily represent the "true" origins or destinations of the air trips. Although 141 zones do not contain CAB air hubs, there are clearly intercity air trips that have these zones as their origins or destinations. Furthermore, passengers, whose origins and destinations are in a given zone, will not always use the airport closest to that zone but may often use a more distant airport offering more convenient services. Various empirical studies (3) support this concept, namely, that the traveler is concerned with the door-to-door service provided by the transportation system and not just with the service on 1 segment of his trip. Therefore, an analysis was developed and carried out to convert the airport-to-airport trip tables into trip tables representing "true" origins and destinations. If A and B represent 2 air hubs, and if $\{i\}$ and $\{j\}$ denote the sets of tributary zones to these hubs, then the problem to be solved was twofold: (a) identify those zones that belong to $\{i\}$ and $\{j\}$, and (b) determine the feasible routes between a given i and a given j that pass through A and B.

The gravity model was used to distribute the users of each air hub to their "true" origins or destinations. Note that, in all likelihood, the sets $\{i\}$ and $\{j\}$ would contain the zone in which A and B were located. The trip-attraction input to the gravity model consisted of the population of each zone. The trip-production input consisted of the total number of trips at each air hub, i.e., the row sums of the airport-to-airport trip tables. As noted earlier, 141 of the 490 zones contained no air hub, and their trip-production factors were set equal to zero. The airport access friction factor curve was based on the results of a recent NCHRP report (4) that investigated airport-access trip production as a function of travel time from and to the airport. A composite curve was developed that incorporated data from airports in Atlanta, Buffalo, Philadelphia,

Pittsburgh, Providence, and Washington, D.C. This curve showed no significant change in trip production for access times ranging from 15 to 60 min.

Examination of the output of the gravity model revealed that about 62,855,000 out of the total of 88,410,000 air trips had a final or true origin and destination that was the same as the zone in which the airport city was located. The remaining 25,555,000 air trips, or about 29 percent, had a final origin and destination that was other than the zone in which the airport city was located. In other words, approximately 29 percent of the total airport access travel in the country crossed a zonal boundary in the 490-zone structure. These results are plausible, despite a tendency for the gravity model to distribute trips to points more distant than might be desired.

After the gravity model had been run, it can be seen that 2 trip tables were created: (a) an airport-to-airport trip table, and (b) an airport-to-true-origin-and-destination-zone trip table. In other words, information was available on the trips from i to A, from A to B, and from B to j. The next step was to develop a trip table for the true origins and destinations of the air trips, that is, to develop an i-to-j trip table. The number of trips from i to j was defined as

$$T_{ij} = \sum_{A,B \in r} \left(\frac{T_{iA}}{\sum_i T_{iA}} \frac{T_{jB}}{\sum_j T_{jB}} T_{AB} \right)$$

where T_{AB} is an entry in the airport-to-airport trip table, T_{iA} and T_{jB} are entries in the airport-to-true-origin-and-destination-zone trip table, and r is the set of all feasible routes between i and j.

Because of the large number of combinations that had to be considered, application of the preceding relationship resulted in high computer costs. To reduce running times to more acceptable levels required that conditions be formulated so that only the more significant entries from each trip table were introduced into the analysis. The first condition, relating to the airport-to-zone trip table, restricted the number of trips on a given interchange to those greater than or equal to 15 percent of the total trips at the given airport. All zone-to-airport interchanges that did not meet this criterion were discarded, and those interchanges that satisfied the criterion were factored so that the total number of trips at the airport remained unchanged. Imposing this condition counteracted the tendency of the gravity model to distribute trips too widely. This criterion also implied that no more than 6 zones could be the true origin and destination for a given airport, which was acceptable in view of the relatively large size of the zones.

The second condition, relating to the airport-to-airport trip table, was that a given interchange had to pass one of the following tests to be considered feasible: (a) It had to be greater than 0.1 percent of the total airport use at both airports, or (b) it had to be greater than or equal to 10 trips. All possible routes between a given i and a given j were checked for feasibility; in particular, for each i-j, the following were checked: i-A-B-j and i-B-A-j. Unlike the application of the former condition, the application of the latter did not make it possible to estimate total trips and to appropriately factor the feasible interchanges to a control total prior to the completion of the analysis.

With the imposition of the conditions identified in the preceding, it was possible to develop the true origin-to-destination air trip tables. A total of approximately 85,823,000 air trips were contained in this table, as compared to a CAB control total of about 88,435,000. Thus, about 2.9 percent of the total air trips were lost through the application of the second condition. Approximately 116 of the 490 zones did not have any air trips originating in or destined to the zones, as compared to a total of 141 zones in the same category prior to the application of the gravity model, which was a reduction of about 18 percent in the number of zones in this category.

DEVELOPMENT OF INTERCITY MODAL-SPLIT MODEL

For noncorridor movements, competition was assumed to be limited to the automobile, air, and bus modes. Bus travel, which could not be modeled because of insufficient data, was estimated based on factors obtained from the 1967 National Travel Survey. The analysis described in this section was designed to evaluate competition between the automobile and air modes. In the following, W, w, t, and c respectively identify numbers of trips, modal splits, travel times, and travel costs, with the subscripts 1 and 2 referring to air and automobile modes respectively; thus, w_1 denotes number of air trips between 2 zones, whereas w_2 denotes number of automobile trips between 2 zones.

Model Development

It is assumed that a traveler selects a mode by comparing the travel times and the travel costs of both modes. It is suggested that this process be described by using the differences between travel times and travel costs. As may be expected, these variables are highly collinear, primarily because both time and cost are estimated as functions of distance. Collinearity problems are avoided by using a single independent variable that is defined as a linear combination of the differences between automobile time and air time on the one hand and automobile cost and air cost on the other. Mathematically, this may be expressed as

$$x = \alpha \Delta t + \beta \Delta c$$

$$\Delta t = t_2 - t_1$$

$$\Delta c = c_2 - c_1$$

where α and β are 2 specified coefficients.

The 2-mode model is based on the hypothesis that a differential change in the share of one mode, such as air, is proportional to the share of each mode and the differential change in the independent variable x. Mathematically, this can be written as

$$dw_1 = pw_1 w_2 dx \tag{1}$$

where p is the proportionality coefficient to be determined by calibration. Between w_1 and w_2, the following relationship $w_1 + w_2 = 1$ must hold. If $1 - w_1$ is substituted for w_2, Eq. 1 becomes

$$\frac{dw_1}{w_1(1 - w_1)} = pdx \tag{2}$$

The integration of this differential equation yields

$$w_1 = \frac{1}{1 + \exp(-px - a)} \tag{3}$$

where a is a constant of integration to be determined by calibration. This function is represented graphically by a logistic curve.

Conversely, considering the differential change in the share of the automobile mode leads to

$$w_2 = \frac{1}{1 + \exp(-qx - b)} \tag{4}$$

where q and b are parameters corresponding to p and a respectively. These 4 parameters must satisfy the following identity:

$$w_1 + w_2 = \frac{1}{1 + \exp(-px - a)} + \frac{1}{1 + \exp(-qx - b)} \equiv 1$$

which implies $\exp[-(p+q)x-(a+b)] = 1$ which holds if $(p+q)x+(a+b) = 0$ for any value of x; that is, if $q = -p$ and $b = -a$. Hence, for calibration purposes, it is sufficient to calibrate either Eq. 3 or Eq. 4.

The variable x measures the difference between automobile and air; hence, an increase in x implies an increase in air trips and a corresponding decrease in automobile trips. This relationship holds only if p is positive, which implies that negative calibrated values of p must be rejected even if other elements of the calibration are satisfactory.

Sensitivity

Sensitivity is defined as the change in modal split due to a unit change in x. If "small" changes are assumed, the differential expression given by Eq. 1 can be used; i.e., in the case of air modal split

$$dw_1 = pw_1(1 - w_1)\,dx.$$

The rate of change in modal split is, therefore,

$$\frac{dw_1}{dx} = pw_1(1 - w_1)$$

The graph of this function varying between 0 and p/4 is represented by a parabola as shown in Figure 1.

The relative change in modal split for a change dx in x is given by

$$\frac{1}{w_1}\frac{dw_1}{dx} = p(1 - w_1)$$

The graph of this function is shown in Figure 1 for the same values of p as used in the preceding graph.

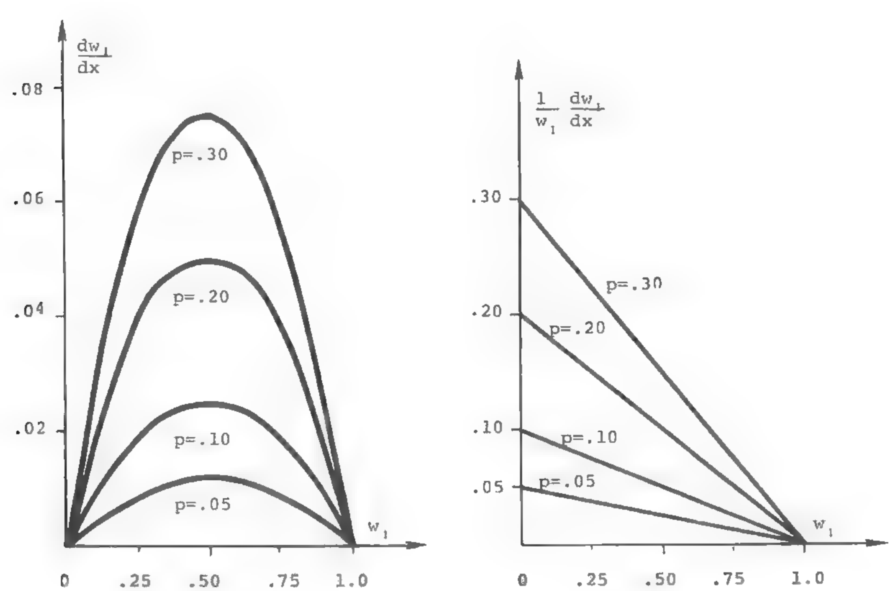


Figure 1. Parametrical sensitivity curves of the diversion model.

As an example, if $dx = 6$, which corresponds to an improvement of air travel time by 1/2 hour (as it will be seen in the retained model), and if $p - 0.02$, the values of dw_1 and dw_1/w_1, expressed in percentages, are as follows:

w_1	dw_1	dw_1/w_1
0	0	12
10	1.08	10.8
20	1.92	9.6
30	2.52	8.4
40	2.88	7.2
50	3.00	6.0
60	2.88	4.8
70	2.52	3.6
80	1.92	2.4
90	1.08	1.2
100	0	0

The change in modal split given by Eq. 1 implies only an aggregate change in x. However, if it is desired to evaluate the variations of w_1 because of variations of a p or x or both that enter in the expression of the modal split, the total differential of w must be used. In such a case, the change in w is

$$dw_1 = \frac{\partial w_1}{\partial x} dx + \frac{\partial w_1}{\partial p} dp + \frac{\partial w_1}{\partial a} da$$

Note that the preceding expressions contain implicitly the components of x (i.e., Δc, Δt, α, and β).

Calibration

The model was calibrated by simple linear regression. For this, Eq. 3 can be written as

$$1 - \frac{w_2}{w_1 + w_2} = \frac{1}{1 + \exp(-px - a)}$$

which is linearized by taking the natural logarithm of the reciprocal of each side, i.e.,

$$\text{Log} \frac{w_1}{w_2} = px + a$$

In the expression of x, the coefficient α was set equal to 0 or 1 whereas β was searched for (by increments of 0.5 starting at 0) to obtain the best possible fit.

Because of the availability of air and highway network data for 1967, base-year air, automobile, and total person trip tables were developed for 1967, which was chosen for calibration of the intercity air-versus-automobile modal-split model. Because a majoı purpose of the study was to analyze the competition between the air and automobile modes and because over 96 percent of the air travel in the United States took place among 144 SMSA's or air hubs, it was decided that the interchanges used to calibrate the air-versus-automobile modal-split model should be selected from a sample of these 144 air hubs.

Two calibration data sets were developed, differing only with respect to their automobile impedances that included or excluded overnight times and costs. Each of the 2 data samples was stratified into 7 categories of trip lengths. Finally, to properly reflect the relative importance of a data point required that observations in each calibration data set be weighted according to the total number of trips relative to a given interchange.

TABLE 1

CALIBRATION RESULTS FOR TRIP LENGTHS GREATER THAN 99 MILES

Independent Variable	p	Standard Error of p	a	R
Δc	−0.0604	0.0010	2.697	−0.66
Δt	+0.1983	0.0029	2.780	+0.68
$\Delta c + 1.0\Delta t$	−0.0688	0.0014	2.244	−0.57
$\Delta c + 2.0\Delta t$	−0.0547	0.0019	1.331	−0.37
$\Delta c + 5.0\Delta t$	+0.0410	0.0012	1.461	+0.43
$\Delta c + 10\Delta t$	+0.0234	0.0004	2.371	+0.61
$\Delta c + 12\Delta t$	+0.0192	0.0003	2.470	+0.63

Results of Calibrations

Observations including automobile overnight cost and time yielded generally poor results. The value of p was not positive for all trip-length strata; even when p was positive, the corresponding correlation coefficients were much too low to be acceptable. When automobile overnight costs and times were excluded, regression runs performed on stratified samples yielded acceptable results for the short-length trips (0 to 99 miles). However, for the other strata as well as the unstratified sample, the same difficulties as in the preceding cases were encountered.

To partially remove the "noise" in the data, we decided to exclude observations relative to trips under 100 miles. The results of selected runs are given in Table 1. The highest correlation coefficient (among those associated with positive values of p) corresponded to $\beta = 0$; i.e., the difference between automobile and air times was the independent variable. Note that, because of the high value of p, the model would be extremely sensitive to time. However, if the independent variable were the difference in cost, p would be negative, and the equation should be rejected. To incorporate time and cost in the equation, together with the correct sign for p and an acceptable regression coefficient, required that α be increased to the vicinity of 10, as given in Table 1. The values of R increased rapidly and stabilized around 0.625 when α is more than 10.

The equation corresponding to $\alpha = 12$ was selected because it gave an acceptable sensitivity. However, the large standard error of the equations must be noted (2.54 for a mean log w_1/w_2 of 0.65, a result of the dispersion of the data). The selected equation (for trip lengths greater than 99 miles) is

$$w_1 = \frac{1}{1 + \exp(-0.0192x + 2.470)}$$

where $x = \Delta c + 12\Delta t$. The graph of the equation is given in Figure 2 (times and costs are in hours and dollars respectively).

The model was tested by comparing estimated trips to observed trips for each interchange. The coefficient of correlation between estimated and observed air trips is about 0.80; that is, about 64 percent of the variance in the actual number of air trips was explained by the estimating equation, a considerable improvement over the coefficient of correlation of the logarithmic equation, which is only 0.625 (about 40 percent of the variance explained). Between zones that were less than or 99 miles apart, a constant 100 percent of the interchange was automatically assigned to the automobile mode.

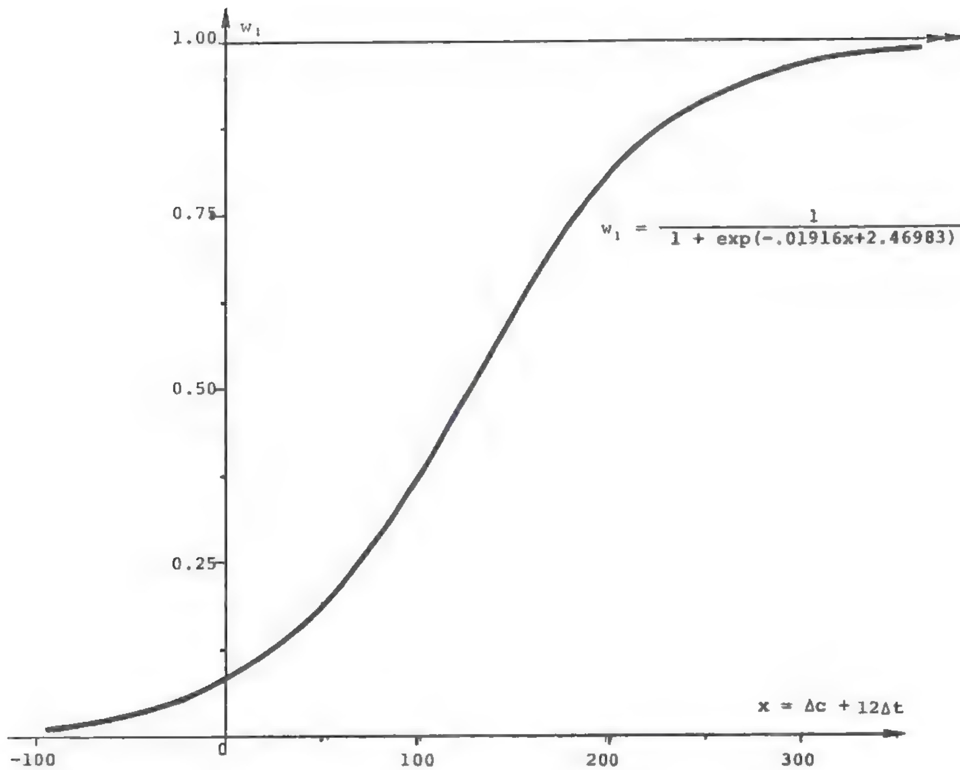


Figure 2. Selected diversion curve of the proportion of air trips.

RECOMMENDATIONS

Several recommendations based on the experience acquired in this study are made for further work. These observations should in no sense be viewed as a comprehensive approach to a demand analysis of national travel.

Results of this investigation emphasize the importance of developing appropriately disaggregate and accurate input data for any analysis of intercity travel demands. The availability and quality of data have been and, in all likelihood, will always be a constraint on the level and quality of an analysis. However, this should not prevent the analyst from developing the best possible solution within the given resources. In view of this fact, any future project to analyze intercity travel demands should carefully incorporate data considerations within the overall study framework, although it clearly would be a mistake if the entire project were limited to data acquisition. Of the 3 types of data required to perform a transportation analysis (namely, activity, network characteristics, and travel), existing travel-pattern data, particularly by automobile, are probably the most difficult to acquire. This analysis suggests the requirement for data on national travel stratified by origin and destination, mode, purpose, and income of the travelers. Other stratification variables might include group size, duration, peak versus off-peak, and occupation.

Additional variables, particularly trip purpose, group size, and income, should be introduced into the demand analysis. Proper use of these variables could, for example, provide a rationale for the parameters of the modal-split model to change through time, instead of imposing the parameters calibrated for 1967 on analyses performed for 1975, 1980, and 1990.

Certain analytical techniques should be introduced into a framework for an intercity travel demand analysis. The access time for air travel constitutes a significant proportion of the total travel time for air, particularly in the 200- to 700-mile distance range in which automobile travel is particularly competitive with air travel. The Access Characteristics Estimation System (ACCESS) that was recently developed and implemented (5) provides a computer-based analytical methodology that can be used to accurately estimate access characteristics for intercity person movements without the relatively expensive acquisition and processing of large amounts of transportation network data.

Although the CAB data probably provide the best available inventory of national travel for a given mode, this information pertains to airport-to-airport travel and not to the true origin and destination of the trip. The technique proposed in this report provides an efficient and potentially reliable method for converting airport-to-airport trip tables into true air origin-and-destination trip tables.

ACKNOWLEDGMENTS

This research was performed for the U.S. Department of Transportation. The authors gratefully acknowledge the assistance of the staff of the Department's Office of Systems Requirements, Plans and Information, particularly that of Carl N. Swerdloff, Arrigo Mongini, and Daniel P. Maxfield. The authors also express their appreciation and thanks to Robert W. Whitaker and Robert L. November of Peat, Marwick, Mitchell and Company, who designed and implemented the computer system.

REFERENCES

1. National Intercity Travel: Development and Implementation of a Demand Forecasting Framework. Peat, Marwick, Mitchell and Company, March 1970.
2. Domestic Origin-Destination Survey of Airline Passenger Traffic, Civil Aeronautics Board, Vol. VIII-13, 1967.
3. Wiggers, G. Interim Technical Report—A Preliminary Report on the Cleveland Before and After Study. Office of the Secretary of Transportation, U.S. Dept. of Transportation, 1969.
4. Keefer, L. E. Urban Travel Patterns for Airports, Shopping Centers, and Industrial Plants. NCHRP Rept. 24, 1966.
5. Access Characteristics Estimation System. Peat, Marwick, Mitchell and Company, Dec. 1969.

THE NATIONAL ACADEMY OF SCIENCES is a private, honorary organization of more than 700 scientists and engineers elected on the basis of outstanding contributions to knowledge. Established by a Congressional Act of Incorporation signed by Abraham Lincoln on March 3, 1863, and supported by private and public funds, the Academy works to further science and its use for the general welfare by bringing together the most qualified individuals to deal with scientific and technological problems of broad significance.

Under the terms of its Congressional charter, the Academy is also called upon to act as an official—yet independent—adviser to the Federal Government in any matter of science and technology. This provision accounts for the close ties that have always existed between the Academy and the Government, although the Academy is not a governmental agency and its activities are not limited to those on behalf of the Government.

The NATIONAL ACADEMY OF ENGINEERING was established on December 5, 1964. On that date the Council of the National Academy of Sciences, under the authority of its Act of Incorporation, adopted Articles of Organization bringing the National Academy of Engineering into being, independent and autonomous in its organization and the election of its members, and closely coordinated with the National Academy of Sciences in its advisory activities. The two Academies join in the furtherance of science and engineering and share the responsibility of advising the Federal Government, upon request, on any subject of science or technology.

The NATIONAL RESEARCH COUNCIL was organized as an agency of the National Academy of Sciences in 1916, at the request of President Wilson, to enable the broad community of U.S. scientists and engineers to associate their efforts with the limited membership of the Academy in service to science and the nation. Its members, who receive their appointments from the President of the National Academy of Sciences, are drawn from academic, industrial, and government organizations throughout the country. The National Research Council serves both Academies in the discharge of their responsibilities.

Supported by private and public contributions, grants, and contracts, and voluntary contributions of time and effort by several thousand of the nation's leading scientists and engineers, the Academies and their Research Council thus work to serve the national interest, to foster the sound development of science and engineering, and to promote their effective application for the benefit of society.

The DIVISION OF ENGINEERING is one of the eight major Divisions into which the National Research Council is organized for the conduct of its work. Its membership includes representatives of the nation's leading technical societies as well as a number of members-at-large. Its Chairman is appointed by the Council of the Academy of Sciences upon nomination by the Council of the Academy of Engineering.

The HIGHWAY RESEARCH BOARD, an agency of the Division of Engineering, was established November 11, 1920, as a cooperative organization of the highway technologists of America operating under the auspices of the National Research Council and with the support of the several highway departments, the Bureau of Public Roads, and many other organizations interested in the development of transportation. The purpose of the Board is to advance knowledge concerning the nature and performance of transportation systems, through the stimulation of research and dissemination of information derived therefrom.

HIGHWAY RESEARCH BOARD
NATIONAL ACADEMY OF SCIENCES—NATIONAL RESEARCH COUNCIL
2101 Constitution Avenue Washington, D. C. 20418

ADDRESS CORRECTION REQUESTED